Climate Modelling User Group (CMUG)

Deliverable 1.1

Meeting the needs of the Climate Community – Requirements

Centres providing input: Met Office, DLR, ECMWF, IPSL, MétéoFrance, MPI-M, SMHI, BSC, CMCC, DMI, NCEO (University of Leicester), NCEO (University of Edinburgh)

Version	Date	Status
0.5	8 July 2019	Input from online survey
0.6	17 July 2019	Input to other sections
0.7	8 August 2019	Input from MétéoFrance, MPI-M, Met Office, SMHI, IPSL, DLR
1.0	29 August 2019	Input from ECMWF
1.1	20 December 2019	To address ESA RID comments. Revisions from Met Office, ECMWF, IPSL, Meteo France
1.2	06 July 2020	Input from David Ford on OC ECV
1.3	10 July 2020	Input from Pablo Ortega on Sea Ice ECV
1.3.1	14 July 2020	Input from Jeff Ridley on Ice Sheets ECVs
1.3.2	27 July 2020	Input from Enza di Tomaso on Land Cover
1.3.3	27 July 2020	Input from Jeff Ridley on Ice mass balance (5.2)
1.3.4	28 July 2020	Input from Debbie Hemming on AGB
1.3.5	28 July 2020	Input from Eleanor Burke on Permafrost
1.3.6	31 July 2020	Input from Jean-Christophe Calvet on snow
1.3.7	04 Aug 2020	Input from Andy Shepherd from IMBIE
1.3.8	10 Aug 2020	Input from Pablo Ortega and David Ford on Sea surface salinity
1.3.9	08 Sept 2020	Input from Axel Lauer on model development and evaluation (3.3) and Water Vapour (4.1)

1.3.10	08 Sept 2020	Input from Angela Benedetti on Requirements for Climate Service Datasets
1.3.11	29 Sept 2020	Input from Grace Redmond and Erasmo Buonomo on Lakes
1.3.12	01 Oct 2020	Input from Gabriel Rooney on Lakes
1.3.13	09 Oct 2020	Input from Pablo Ortega on SST and Clouds
1.3.14	12 Oct 2020	Input from David Ford on SST
2.0	23 Oct 2020	Final edits by MO and submit to ESA
2.0.1	02 Nov 2020	Input from Axel Lauer on Clouds and Updates in response to ESA RID for v2.0
2.0.2	09 Nov 2020	Input from Angela Benedetti on Water Vapour, in response to ESA RID for v2.0
2.1	09 Nov 2020	MO response to all other comments in RID for v2.0 and resubmit to ESA
2.2	11 Jan 2021	Comments from ESA addressed, and final version 2.2 created
3.0	24 Jun 2022	New version created for Milestone 6
3.0.1	27 Jun 2022	Input from Erasmo Buonomo on Lakes
3.0.2	27 Jun 2022	Input from David Ford on Sea State
	27 Jun 2022	Check of Sea Level, OC and SST by David Ford
3.0.3	27 Jun 2022	Input from Pablo Ortega on Sea Surface Salinity
	28 Jun 2022	Check of cloud section by Axel Lauer
3.0.4	30 Jun 2022	Check of Lakes section by Grace Redmond
3.0.5	02 Jul 2022	Input from Andreas Wernecke on Ice Sheets and SI
3.0.6	15 Jul 2022	Input from Angela Benedetti on WV, O3, GHG and AER
3.0.7	18 Jul 2022 18 Jul 2022	Input from Pablo Ortega on Fire Input from Catherine Ottlé on Snow and Land cover
3.0.8	22 Jul 2022	Check of LST section by Rob King
3.0.9	25 Jul 2022	Check of Permafrost section by Jean-Christophe Calvet
3.0.10	26 Jul 2022	Check of Biomass and Soil Moisture by Daniele Peano
3.0.11	05 Aug 2022	Final edits and submit to ESA
3.1.0	08 Mar 2022	Input on vegetation and river discharge from Daniele Peano
3.1.1	28 Mar 2022	Input on vegetation requirements from Souhail Boussetta at ECMWF
4.0	31 May 2024	First version with new CCI ECVs added submitted to ESA

Deliverable 1.1

Meeting the needs of the Climate Community – Requirements

Contents

Meeting the needs of the Climate Community – Requirements Baseline

1. Purpose, scope and construction of the Requirements Baseline

The purpose of this document is to assist the CCI ECV projects and cross-ECV demonstration projects in focusing on the needs of the Climate Modelling Community (CMC), Climate Research Community (CRC) and other expert users of climate data. It aims to do the following:

1) present an analysis of the satellite climate observation data requirements of the CMC, CRC and other expert users of climate data.

2) cover both the requirements for the 23 ECVs in terms of parameters, resolution and errors/uncertainties and also, where appropriate, cover the requirement for observation operators for each of the ECVs.

3) address the requirements for CCI datasets to be included in the Copernicus Climate Change Service (C3S) and the obs4MIPs interface.

4) cover overarching technical requirements and scientific linkages for the datasets produced.

This document confirms and builds upon the user requirements inventoried by CMUG in Phase 2 of the CCI programme (~ 2014–2017)[1]. The new information found here is of greater detail in describing user needs, from an extended base of users interviewed and from users experienced in using CCI data. It also adds sections on the 9 new ECVs from CCI+. It is acknowledged that the climate data needs of the climate research community are evolving. CMUG, through its interactions with this community, will ensure knowledge of user requirements is up to date and relevant.

A key example of user requirements across the CMC applies to the obs4MIPs[2] initiative (Teixeira *et al*., 2014) that provides an archive of gridded Earth climate system observations to facilitate model evaluation in the CMIP6 initiative (Meehl *et al*., 2014). This is the gold standard of climate observational data sets and the aspiration is to have all ESA CCI ECV datasets included in obs4MIPs. To achieve this individual CCI datasets conforming to strict data and meta-data formats required by obs4MIPs and supporting technical notes need to be created. CMUG (through WP5) then facilitates the publishing of these datasets and associated metadata on the obs4MIPs database. The CCI data submitted to obs4MIPs will sit alongside other observational data used for model evaluation and will be implemented in ESMValTool[3] by

[1] CMUG Phase 2 Deliverable 1.1: User Requirement Document (v0.6), available at http://ensembles-eu.metoffice.com/cmug/CMUG_PHASE_2_D1.1_Requirements_v0.6.pdf.
[2] https://esgf-node.llnl.gov/projects/obs4mips/
[3] https://github.com/ESMValGroup/ESMValTool

CMUG to routinely benchmark the models against ESA CCI data. CMUG WP5 focuses on applying CCI data to evaluate climate models using ESMValTool.

The tables provided for each ECV provide details of user requirements from a number of sources and where applicable follow the Goal/Breakthrough/Threshold approach adopted in the GCOS ECVs Requirements document (GCOS, 2022):

- Goal (G): an ideal requirement above which further improvements are not necessary.
- Breakthrough (B): an intermediate level between threshold and goal which, if achieved, would result in a significant improvement for the targeted application. The breakthrough value may also indicate the level at which specified uses within climate monitoring become possible. It may be appropriate to have different breakthrough values for different uses.
- Threshold (T): the minimum requirement to be met to ensure that data are useful

2. Introduction

Climate researchers require ongoing global and regional measurements of the Earth's climate to monitor climate variations on all scales and to evaluate or provide input into climate models. Given their global and temporal coverage and spatial resolution satellite data are well placed to meet this need. For some variables satellite datasets now span a period of more than 35 years, which is a long enough time series for use in climate monitoring, model initialisation and model evaluation, provided certain requirements can be met.

The uncertainty characteristics of the satellite measurements must be understood and quantified; otherwise, little confidence can be placed in the derived climate data records (CDRs). Because most of the measurements were not taken with climate applications in mind, the data need careful preparation for climate monitoring. Also, satellites do not make localised 'conventional' *in situ* measurements of e.g. temperature or moisture as represented by climate models, but measurements of indirect parameters e.g. upwelling radiance or GPS signal refraction angles. For some parameters, climate models can deal with this by including 'observation simulators' or 'forward operators' to compute the variable measured by the satellite from the model fields, thus avoiding some of the uncertainties in the retrieval of conventional variables from satellite data. However, it is important that these simulations can be interpreted in terms of standard geophysical variables, or physical properties such as humidity, cloud drop size or crystal shape, as model parameterisations are often framed in terms of these physical quantities. It is also important that the error characteristics of the observation simulators are well understood and documented.

Climate researchers usually confront models with observations with the following aims:

- To interpret the observations and explain the causes of observed variability and change
- To evaluate, constrain and improve climate models, thus gaining confidence in their projections of future change
- To initialise models for reanalyses, seasonal and decadal timescale predictability (data assimilation) and to provide representative initial conditions for climate model simulations

- To prescribe boundary conditions of quantities that are not prognostic variables in climate models

Accordingly, the generic requirements for satellite data are:

- to provide long term monitoring datasets of particular parameters with or without *in situ* data to ascertain decadal and longer-term changes. Models can then be used to attribute the observed variations to natural and anthropogenic forcings and internal variability (IPCC WG1 AR5, 2013).
- to provide long term sets of validated, high quality climate data, with good uncertainty characterisation and documentation for Earth system model evaluation.
- to compare measured parameters, or combinations of observed and/or reanalysed parameters, with model equivalents on hourly up to decadal timescales, to assess the processes and biases in the models and if necessary, to constrain, the processes.
- to initialise seasonal forecasting models, for example with realistic estimates of soil moisture and sea surface temperature.
- to help evaluate the skill of seasonal to decadal forecasts.
- to interpret short term variations of the climate in the long-term context, as in the recent hiatus in observed surface warming.
- to help identify biases in the current and past *in situ* observing network. E.g. comparisons of Microwave Sounding Unit (MSU) retrievals to "families" of radiosondes for identifying shortcomings both in the raw radiosonde data and the satellite datasets.
- to provide homogeneous data, with good estimates of random errors and bias-correction uncertainties, for reanalyses. Existing reanalyses are already very useful for model evaluation, especially in combination with independent satellite data; but the next generation of reanalyses also needs to be sufficiently homogeneous to allow the estimation of long-term trends (Simmons et. al. 2014). In addition, especially in areas with sparse sampling like the polar regions, different reanalysis products differ significantly from one another.
- to provide long term sets of validated, high quality climate data, with good uncertainty characterisation for use by climate service providers.

Now that many satellite climate data records are reaching 35 years in length, they have become an important source of data for use in climate research and the CMC and CRC need to make best strategic use of the opportunities provided by satellite data. Only after quality assurance is demonstrated, can high quality climate datasets be produced that are fit for onward use in an operational or wider societal application. Opportunities for exploitation of the CCI datasets now exist in various activities related to climate services in both national and international arena. The improved interface to climate modellers provided by the ESA Open Data Portal[4] and obs4MIPs project are other channels supporting the uptake of CCI data. Providing CCI data to these interfaces imposes certain requirements on the datasets which are given in sections 6 and 7.

The requirements outlined in this document were captured by CMUG through feedback from >50 experts. The responses given by climate modellers are representative of the full range of

[4] http://cci.esa.int/data

models and the applications operated by them. Experts in the wider CRC responded with information from the areas of climate services (including the Copernicus Climate Change Services and national Climate Service Centres), detection and attribution of climate change, climate process studies, climate change impacts and climate/environmental monitoring. Other groups of users such as the MIP projects, CORDEX and the EEA were also consulted, including many H2020 projects (CRECP, APPLICATE, PRIMAVERA, CRESCENDO), as well as the precursor projects for the Copernicus Sectoral Information Services, the work for which is being conducted by research institutes that are CMUG partners. The contributions from each of these sources is integrated into the sections in this report on each relevant ECV.

Section 3 identifies in more detail those generic application areas where satellite datasets are required for climate modelling. Section 4 outlines the specific requirements for the satellite climate data records (CDR) for the 23 CCI ECVs, where the requirements are described in detail for the 9 new ECVs in CCI+ and refreshed for the 14 ECVs established in CCI Phases 1 and 2. Section 5 lists cross-ECV requirements, including the Earth system budget closure projects sponsored by ESA. Sections 6 and 7 cover the requirements for climate services and obs4MIPs. Section 8 lists the requirements for other ECVs. Section 9 gives the requirements for observation simulators and other tools required by climate modellers to exploit the datasets. Section 10 outlines the technical requirements for data formats, projection, access, etc. Section 11 summarises the key points of this report. Section 12 lists the references. A list of acronyms and definitions of various terms is in section 13. Finally Annex 1 summarises results of the CMUG survey of the CMC and CRC carried out in early 2019.

3. Generic requirements for climate applications

Table 1 summarises the generic requirements for 23 ESA CCI ECVs from a survey (May to July 2019) of experts from climate research centres that was conducted by CMUG. All application areas are mentioned but the comparison with models for model evaluation and development is the most frequently reported use (See Annex 1). It should be noted that the high number of experts who are using, or intending to use, CCI datasets for model development and validation is well served by the ongoing CMUG work with ESMValTool. The rest of this section summarises these results and describes the context to the climate applications in Table 1.

CCI ECV	Model Initial-isation	Prescribe Boundary Conditions	Re-analyses	Data Assimilation	Model Development and Validation	Climate Monitoring/ Attribution	Q/C in situ data	Climate process study	Other, inc. Climate Services
Atmospheric									
Water Vapour	X	X	X		X	X		X	X
Clouds	X	X				X	X	X	X
Ozone	X	X	X	X	X	X	X		X
Greenhouse Gases	X	X	X	X	X	X	X		X
Aerosols	X	X	X	X	X	X		X	X
Oceanic									

Sea State	X	X	X	X				X	X
Sea Surface Salinity	X	X						X	X
SST	X	X	X		X	X			X
Sea Level	X	X	X	X	X	X			X
Sea-ice	X	X	X		X	X		X	X
Ocean Colour				X	X	X		X	X
Terrestrial									
Above Ground Biomass	X	X		X	X	X	X	X	X
Land Surface Temperature	X	X	X	X	X	X	X	X	X
Permafrost	X	X	X		X	X		X	X
Lakes	X	X			X		X	X	X
Glaciers	X	X			X	X		X	X
Snow	X	X			X	X		X	X
Land Cover	X	X	X		X	X	X		X
Hi-res Land Cover	X	X				X		X	X
Greenland Ice Sheet	X	X			X	X		X	X
Antarctica Ice Sheet	X	X			X	X		X	X
Fire / Burnt Area	X	X		X	X	X	X	X	X
Soil Moisture	X	X	X	X	X	X	X	X	X
Anthropogenic									
Total water use									

Table 1. Use of CCI ECVs for different climate applications. This table updates results from previous CMUG reports and covers known research and other applications for all CCI ECVs, gained from the CMUG user surveys and expert interviews. FAPAR and LAI which will be included in the next CCI+ phase for the Vegetation ECV project are not addressed in this report.

3.1 Climate monitoring and attribution

Satellite datasets need to span at least several decades in order to meaningfully monitor climate change. Some satellite datasets already approach 35 years in length, but many are shorter than 20 years although continually expanding.

Climate monitoring imposes the most stringent requirements for satellite data both in terms of stability of the measurement and the minimum time period of the dataset. In addition, significant overlap periods between successive sensors are required to ensure the fidelity of the time series, as recommended by the GCOS monitoring principles[5], in particular item 12.

Time series of greenhouse gas, ozone and aerosol concentration profiles and total column amounts are important for trend analyses as significant increases or decreases will affect the

[5] http://ane4bf-datap1.s3-eu-west-1.amazonaws.com/wmocms/s3fs-public/ckeditor/files/GCOS_Climate_Monitoring_Principles.pdf?l1e4ALNYxVIStmm19we2Sz0evxEFpHmT accessed 06-10-2020.

atmospheric radiative balance. Global coverage allows regional and/or temporal variations to be investigated and potentially attributed to natural or anthropogenic causes. Changes over the long term in cloud amount and water vapour are also associated with atmospheric feedbacks due to climate change and are examined in attribution research.

The ocean ECVs of sea level, sea-ice coverage and thickness are critical parameters that must be monitored as key indicators of climate change. Sea surface temperature (SST) is an ECV which has been monitored by in-situ observations since the mid 1800's and so is an excellent indicator of climate change. The complication with satellite measurements of SST is that they measure the skin not the bulk SST and so a "correction" has to be made to the satellite CDRs of SST to obtain a "bulk" SST as would be measured by ships and buoys. The record for ocean colour measurements is relatively short but is now available as a 22-year long dataset on the CCI Open Data Portal and having passed the 20-year mark becomes more useful for long term climate studies. The new CCI+ marine ECVs of Sea State and Sea Surface Salinity are of interest in other areas of climate research such as initialising and evaluating models.

For the land surface, land-cover, biomass and fires are important to help monitor and understand the carbon cycle. Records of decadal changes in land-cover, above ground biomass (AGB), fire numbers and burnt area help to show the amount of deforestation occurring in the last two decades, which is an indicator mostly of human activity but also of environmental and climate change impacts. The extent of permafrost, snow, ice sheets, glaciers and ice caps is also an important indicator of climate change and satellite data can complement the ground-based observations. Long term monitoring of these variables provides evidence of changes in the climate system that are annually variable but display a long-term trend. The irreversibility of changes in some of these cryosphere ECVs makes them key to understanding the climate system and future possible changes in it. The run-off from melting glaciers and the Greenland ice sheet over the long term are also of interest to climate change attribution studies and are linked to other climate change impacts such as sea level rise and on the strength of the North Atlantic meridional overturning circulation.

An area of interest in climate monitoring is the assessment of rapid/irreversible climate changes which requires confidence in the prediction of the thermohaline circulation and carbon cycle/sea ice non-linear feedbacks. Close monitoring of greenhouse gas concentrations and sea-ice coverage/thickness from satellites is important to provide early warning of any sudden changes. Fire, soil moisture and vegetation are also examples of variables that can change rapidly and have significant impacts, for example rapid die-back in the Amazon.

Finally, there are some satellite derived metrics, which are not ECVs as defined by GCOS, but nevertheless are of interest. Severe weather events such as the annual number of tropical cyclones in each ocean basin, frequency of intense extra tropical storms, severe drought episodes and heat waves are all of interest for climate change and applications studies and can be inferred from satellite data with some effort. There is a need from policy makers and other decision-makers and users for a better understanding of the risk of current extreme weather events and the extent to which this risk has changed as a result of human influence. Some of the ESA ECVs may contribute to these metrics and the requirements need to reflect this.

The requirements for climate monitoring measurements are stringent. For example, an SST decadal trend of 0.2 °C per decade requires the satellite CDRs to have a stability of <0.05 K (it is important to distinguish between stability and accuracy here). For climate trends the measurements must be stable over long time periods and any changes must be understood and accurately modelled. Requirements on the bias can be less stringent so long as there are other complementary measurements to compare with. The GSICS project is putting in place an infrastructure to provide these measurements to estimate and monitor biases in Level-1 data from different sensors. Therefore, one of the requirements on the ECVs that users indicated would be useful for climate monitoring and attribution is that they make use of the GSICS[6] measurements to ensure their accuracy can be traced back to International Standards as addressed by the CEOS QA4EO project[7].

3.2 Model initialisation and definition of boundary conditions

A current major application of satellite data continues to be, together with conventional *in situ* data, defining the initial state of the atmosphere/surface for NWP models and decadal prediction systems. Reanalyses are important examples of this requirement, for example the 5th ECMWF Reanalysis (ERA5) and products derived from it, such as the high resolution land reanalyses ERA5 Land which has a 9 km grid and hourly time step[8], these require long and stable time series for initialisation. The Above Ground Biomass (AGB), Land Surface Temperature (LST) and Permafrost terrestrial ECVs in CCI+ all have strong potential to contribute to future reanalyses products.

For initialisation of 'present-day' coupled climate control experiments the atmospheric state is not as crucial as these models equilibrate to their own climate state independent of the initial state. But it is still preferable to start from realistic initial conditions in order to avoid spending large amounts of computational time on reaching the native equilibrium, and to be able to judge the growth of errors without massive drifts.

Most of the 23 CCI ECVs, including all nine of the new CCI+ ECVs, have potential for model initialisation (see Table 1) primarily through improving the representation of the surface fields. The stability and accuracy requirements for initialisation are more relaxed than for climate monitoring as the initial uncertainties in the model fields without the observations are often far greater than the measurement uncertainty.

Land cover type is an ECV required as a model surface field as it can affect the local radiation and provide sources and sinks of various atmospheric variables (e.g. aerosols, CO_2, CH_4 etc.). All NWP and climate models use land cover to define the surface types in or initialise their land surface models. Information on soil moisture dynamics is of major importance as soil moisture has a primary effect on the land surface memory and the partitioning of surface heat and moisture fluxes.

[6] https://gsics.wmo.int/en
[7] Quality Assurance Framework for Earth Observation (QA4EO) [http://www.qa4eo.org/]
[8] https://climate.copernicus.eu/c3s-releases-first-instalment-era5-land-dataset-land-based-studies-and-applications#

3.3 Model Development and Evaluation

Satellite observations are a key part of the development and evaluation of climate models. Banks *et al*. (2008) present assessment criteria for the Hadley Centre model, HadGEM3, where components of HadGEM3 were found to be sensitive to atmospheric and ocean fluxes, e.g. land surface temperature (particularly northern continental summer temperature), rainfall over land (particularly Indian sub-continental rainfall in northern summer), soil moisture, and dust concentrations over both land and ocean (Banks *et al*., 2008).

Coupling the various components in climate modelling is a priority in climate research. For instance, the coupling between atmospheric chemistry (air quality, oxidation, stratosphere-troposphere processes, ozone hole, etc.) and climate. Although the current generation of tropospheric ozone models is generally successful in describing the principal features of the present-day global ozone distribution, there is less confidence in the ability to reproduce the changes in ozone associated with perturbations of emissions or climate, and consistency between the processes described in the models has to be checked. The observations of the various ECVs allow checking of this consistency, and if appropriate can be used to improve the bio-geo-physical-chemical schemes used in the models.

Long term vertically resolved data sets of constituent observations are required to assess Chemistry Climate Models (CCM). This includes ozone, but also other species that are used to diagnose processes involved in CCM: transport, chemistry, radiation, and dynamics. Such observations are required by CCM validation exercises like CCMVal-2 (see overall recommendations in executive summary, SPARC CCMVal (2010)).

CCI ECVs water vapour, clouds, and aerosols are important to validate model fields. For example, the accurate representation of water vapour and cloud properties in climate models is important to reduce the range of uncertainty in climate sensitivity as the response of clouds to climate change is contributing significantly to the high uncertainty in modelled climate sensitivity. Datasets of cloud properties (i.e. fractional cover, top height, phase, microphysical properties etc.) provide an important constraint for climate models and allow validation of cloud parameterizations needed to describe processes that the models cannot resolve explicitly. Regional estimates of all these parameters are also important for detection/attribution studies. In addition, instantaneous estimates of cloud parameters are also important to monitor the diurnal to annual cycles of clouds. In order to compare satellite clouds (e.g. from ISCCP) with model clouds a cloud simulator (sec 9.1) is needed. The Met Office Hadley Centre (MOHC) has developed the CFMIP Observational Simulator Package (COSP) to enable such comparisons[9].

The oceanic ECVs also provide important insight into model quality. For example, some of the longest-standing biases in most large-scale model simulations relate to sea-surface temperature biases in the low-latitude ocean around South America. For better understanding of and eventually reducing these biases, reliable satellite observations of oceanic variables is crucial.

[9] https://www.earthsystemcog.org/projects/cfmip/

Information on the uncertainty in satellite retrievals is also required, something that was not available in such products until the advent of the ESA CCI ECVs.

Terrestrial ECVs contribute to model development and evaluation, for example the observed changes in AGB, LST, Land Cover, Lakes, Permafrost, Snow, Fire and SM contribute to a process level understanding in climate models. Some of these ECVs will inform budget closure projects that have a terrestrial element, such as the RECCAP project on the regional carbon cycle or the International Land Model Benchmarking Project (ILAMB)[10].

ECV accuracy requirements for model evaluation use are driven by the magnitude of the model error as long as the accuracy is better than the model error the dataset is useful. In all cases this requirement will be more relaxed than for use in climate monitoring. Information on the observational uncertainties is helpful in assessing whether climate models are able to reproduce the statistical properties of an observed quantity and to interpret differences between model and observations.

Evaluation of climate models through the systematic application of community tools with agreed benchmark datasets was in common use by the time of the CMIP5 evaluation of climate models for the IPCC AR5 (IPCC, 2013). The current generation of tools for evaluating Earth System Models under the ESMVal initiative are now being employed on model outputs from CMIP6. The evaluation of climate models operated by CMUG partners using CCI datasets and the latest version of the ESMValTool will help to (a) validate the models used by CMUG partners and the climate modelling community (CMIP6), (b) compare CCI data sets used to other observational datasets available, and (c) extend and improve the ESMValTool.

3.4 Input to reanalyses

Global and regional atmospheric, land, and ocean reanalyses are now being undertaken in a number of centres to provide a consistent analysis of the atmosphere over a long time period, typically 40-100 years by using these variables to constrain a Numerical Weather Prediction (NWP) model. Increasingly these reanalysis datasets are being used for climate applications. A key requirement for the data assimilated into these reanalyses is that they are uniformly processed without the discontinuities often seen in operational datasets caused by advances in the real time operational processing of the instrument data.

Accordingly, satellite climate data records are well suited for reanalyses provided they come from a stable processing environment and provide associated error estimates. For the recent ECMWF reanalysis (ERA5) satellite agencies provided homogenous datasets.

In general, reanalysis applications require single-sensor products rather than merged products. Furthermore, these applications often ingest Level-1 satellite data rather than Level-2 retrievals and thus there is a strong interest in uniformly processed Fundamental Climate Data Records (FCDR).

[10] https://www.ilamb.org/

It is worth noting that comprehensive multi-decadal reanalyses are substantial computational projects with demanding production schedules. Uptake of CCI ECV products would be increased if the ECV production timelines can be coordinated with such activities, and CMUG is in a position to keep the ECV projects informed of relevant reanalysis plans.

3.5 Data assimilation for seasonal and decadal forecasts

Recently the need for better initialisation of seasonal and decadal hindcast or forecast models in the operational forecasting centres has become apparent. The oceanic variables with sufficient inertia to act as forcing for seasonal time scales include sea surface temperature, salinity and sea-ice thickness and concentration. Proper initialisation of land surface temperature, soil moisture, snow cover and depth, and aerosol concentration can also increase prediction skill. Vegetation type (Land Cover and AGB) is of interest particularly if coupled with a vegetation model though a good high-resolution dataset of recent vegetation distribution and its conditions (e.g. albedo, leaf area index (LAI)) is valuable in its own right.

Because of its importance to hazardous weather e.g. better monitoring and prediction heat waves, soil moisture is also assimilated in NWP models and used for the initialization of seasonal to decadal climate prediction systems.

Interactions between the polar stratosphere and the mid-latitude troposphere occur on the timescale of a few weeks, and the initialisation of the former can aid the prediction of the latter especially in the first few weeks of seasonal forecasts (Scaife *et al*., 2005). Stratospheric temperature, winds and gas concentrations are therefore of interest to define the model initial state. These parameters can now be measured by satellites to a reasonable degree of accuracy.

The experience of satellite data assimilation at NWP centres, which provides the major impact on forecast skill, can be applied to these longer-range model initialisation problems, in particular from seasonal to decadal forecasts. The atmosphere is represented by at least 80 levels from the surface to 0.1hPa with a horizontal grid size of 10-20km. Only satellite data can provide truly global coverage at this horizontal scale although radiosondes will still have better vertical resolution. In contrast, for reanalyses the satellite climate data records are assimilated in order to affect the short-range forecasts. For a particular ECV to be assimilated it must be represented within the model as a prognostic variable. Table 1, above, shows those CCI variables where data assimilation is currently feasible.

3.6 Climate Services

The Copernicus Climate Change Service (C3S) was initiated in 2014 as part of the EU's Copernicus programme, the environmental and emergency monitoring service for Europe. Its purpose is to provide accurate and independent information for climate security in Europe, although much of its data output is global. C3S is an operational service managed by ECMWF, together with other modelling centres and climate data providers subcontracted to them to provide operational services. In addition to in-situ climate observation data, C3S uses climate quality data produced mainly from satellite observations, and the service will include new ECVs in future as new data comes online (e.g. from future Sentinels).

CMUG engaged previously with precursor projects to C3S. These projects developed different aspects of the service and CMUG gathered information and feedback about their data requirements. The interviews conducted for this requirements analysis included experts currently working on CRECP[11] (Copernicus Roadmap for European Climate Projections) as well as from a broader climate service user community.

The pathway for CCI data to be used in C3S is now clear, as the following CCI data are now available on the C3S Climate Data Store (CDS): Sea Level Rise, SST, Ocean Colour, Sea Ice, GHG, O_3, Aerosols, Clouds, Fire, Ice Sheets, Land Cover, Soil Moisture, and Glaciers. On the condition that new CCI+ data meet required standards they could be included in the CDS and made available to users as early as 2022. CCI ECV data are also used by other Copernicus Services, e.g. Sea Level in the Copernicus Marine Environment Monitoring Service (CMEMS) (in CMEMS the data for sea level rise are produced differently from the sea level rise data in C3S, making it less suitable for use in climate studies), and CCI GHG, O_3 and aerosol ECVs in CAMS. The CCI ECVs currently stored in the C3S CDS are summarised in Table 2.

CCI	CCI+	C3S
	AG-Biomass	
Aerosol		Aerosol
		Albedo
Clouds		Clouds
		Earth Radiation Budget
		fAPAR
Fire		Fire
GHG		GHG
glaciers		Glaciers
Ice sheets		Ice Sheets
Land cover		Land Cover
	HR Land Cover	
		LAI
	Lakes	Lakes
	LST	
Ocean colour		Ocean Colour
Ozone		Ozone
	Permafrost	
		Precipitation
	Salinity	
Sea ice		Sea Ice
Sea Level		Sea Level
	Sea State	
	Snow	
Soil Moisture		Soil Moisture
SST		SST
		Surface Radiation Budget
	Water vapour	

Both in CCI and C3S (blue)
In CCI/CCI+ and not in C3 (red)
In C3S and not in CI/CCI+ (green)

Table 2: CCI ECVs and their availability through C3S (as of July 2019).

The integrity of CCI data is a fundamental requirement to its adoption by C3S, and part of this is the ‘line of sight’ back to documented user requirements (such as this). ‘Climate quality’ of the data is a second requirement for users, which concerns bias correction, accuracy and representation through space and time. The data must also include information about their

[11] https://climate.copernicus.eu/copernicus-roadmap-european-climate-projections

usability and uncertainties and be easily accessible to users. All these quality aspects should be captured in both maturity indices and metadata commentary. The data flows between the main producers and user community are shown in Figure 1. These parallel and serial chains of data processing and application respond to a multi-faceted set of user requirements.

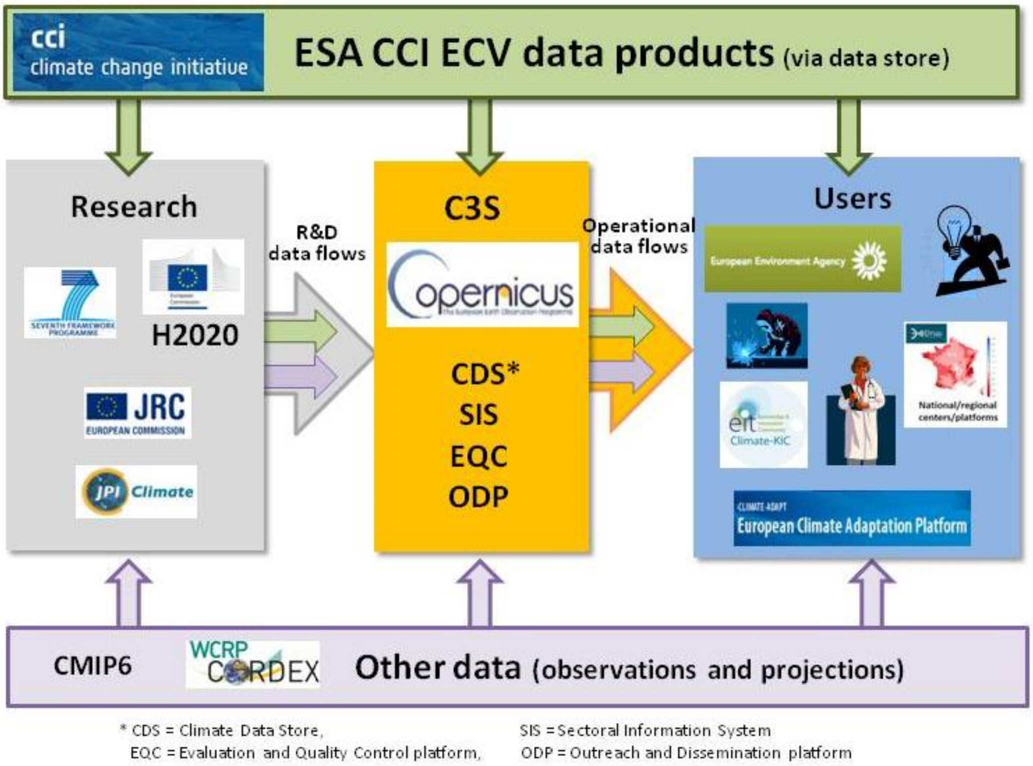


Figure 1: Shows the data flows of CCI data to the research community, C3S and direct to other users, and the subsequent dependencies of downstream users including C3S. It also illustrates the combination of CCI data with data from other sources, for the provision of C3S.

Climate services also exist at a national level and have similar, almost identical, requirements to the C3S but on a finer scale. For example, in the UK a Climate Change Risk Assessment was published in 2017[12] defining user requirements for the UK with respect to addressing climate change on different systems and sectors (e.g. agriculture, industry, health, infrastructure and the natural environment). The report sets out the main priorities for tackling the risk including environmental monitoring and climate services.

In some areas the data requirements for users of climate services are not fully clear and will evolve in future as these services develop. This can be explained with the following example - for monitoring of climate hazards and extreme climate events it is essential that the datasets produced can be used to a) calculate anomalies and b) are available within a short time period (two weeks) after acquisition. A concept similar to reanalysis data would be useful if applied to the satellite data streams, where a consistent data processing is done for a long time period to generate a climatology and process more recent data with the same algorithms to enable anomaly calculations. In parallel, some final datasets with improved algorithms could be

[12] https://www.theccc.org.uk/tackling-climate-change/preparing-for-climate-change/uk-climate-change-risk-assessment-2017/

generated over shorter time spans (of up to a year) as these approaches develop their requirements will evolve and the next version of this document will attempt to capture them.

3.7 Climate Studies (non-modelling)

There are a number of other research areas that use, or have the potential to use, CCI data. These are often specialised, with a core of experts at the centre of the community who often act as both data provider and user. Such communities are:

1. **World glacier inventory** (serviced by the Glacier CCI) which is a unique resource for glacier monitoring and research. Climate modellers are not using this very high-resolution information yet except for regional studies.

2. **Environmental monitoring** - The aerosol, ozone, and GHG ECVs are input to the Copernicus Atmosphere Monitoring Service (CAMS) to provide forecast products of atmospheric composition and air quality in addition to reanalyses of atmospheric composition. The CCI SSH ECV is input to the Copernicus Marine Environment Monitoring Service (CMEMS), although here it is noted that the data are processed and presented in a different way to sea level data on C3S, making them more suited to NWP and less suitable for use in studies of multi-decadal climate.

3. **Many land use studies** – The Land Cover, Soil Moisture and Fire ECVs are in production for the Copernicus Global Land Monitoring Service. The LST, Lakes, and AGB ECVs have the potential to be used by the Copernicus Global Land Monitoring Service to support the provision of products for agriculture, forest, hydrology, etc. to users.

4. **Quality control of in-situ data** - Satellite data can be used to validate in-situ measurements by using the large-scale attributes of the satellite data and assuming any bias is stable over large spatial (>1000 km) and temporal (>1 hr) scales. The requirement is for the stability of the satellite CDR to be more stable than the in-situ measurement errors being validated and so this varies on a case by case basis. If the in-situ measurements are accurate and only have small drifts, then the accuracy (stability and bias) requirements on the satellite data can be high. An example of this might be the use of AATSR brightness temperatures to validate drifting buoy sea surface temperature measurements. The latter can often be in error by several degrees and so an accuracy requirement on AATSR for this application need only be 0.5 K to show useful results. This is a much lower accuracy than the requirement for climate monitoring.

5. **As input data for adaptation research** – adaptation to climate change is an area of research which frequently combines data from the natural world and managed systems with socio-economic data to understand vulnerabilities to and the risks from climate change. CCI data can be of use to this group if the data meets their, often high, spatial and temporal requirements.

6. **As input data for other studies** – there are an increasing number of environmentally focused research areas which combine data on the natural world from different sources to better understand natural phenomena. Such an example is phenology where the timing of

natural events is recorded and analysed. CCI datasets of LST, Land Cover, Fire, Permafrost, AGB, SST, Soil Moisture, Sea Ice and Ocean Colour support this user community.

3.8 Quality Requirements

In addition to the CDR quality requirements described in the above applications, there is a generic set of principles for ECV data quality. Traceability from satellite measurements through bias correction to ECV data is essential for the integrity of any Climate Data Record. The GSICS initiative is therefore key to improving the quality of the global satellite datasets. Another initiative, the now finished QA4ECV[13] project, developed a system for quality controlling ECV datasets so that they have 'climate quality' with respect to both observed long term trends and variability. This project was driven by the user needs of the Climate Services community. It is noted that quality assurance in CCI data production chains is a requirement for production of long-term climate quality data. The FIDUCEO[14] project developed methods and tools for producing fidelity in satellite derived CDRs, these should be used by the CCI ECVs and the standards adhered to in order to produce climate quality data.

[13] http://www.qa4ecv.eu/
[14] http://www.fiduceo.eu/

4. Synthesis of requirements for CCI ECVs

CMUG has undertaken a review of the requirements for the 23 CCI ECVs through direct interactions with expert users and responses to a questionnaire. This report builds on the user requirements published by CMUG in Phase 2[15] and presents an analysis of the user input together with the GCOS requirements (GCOS 2016).

The requirements from the CMC and other expert users given here are in addition to those collected by the CCI projects. An underlying assumption in this requirements definition process is that the CCI datasets produced will be *better* than any existing satellite CDRs and a table of requirements for each ECV is listed in the sub-sections below. Note that it is difficult to be too prescriptive for accuracies as this depends on the horizontal scale chosen to represent the parameters, so for example an SST at a 50 km scale may be more accurate than at a 1 km scale.

An important requirement for all the CDRs is to include their associated uncertainties for each observation where possible and to document the dataset and its uncertainties well. For many applications it is crucial to have a separation of the total measurement uncertainty into random and systematic components. Also, the error correlations between variables are important to consider. A consistent description of uncertainties is needed by users and there are different requirements related to uncertainties in different applications. Descriptions of the different sources and types of uncertainty that were employed in CCI Phase 1 and are also considered here are given in Table 3. The question of how to derive and present uncertainty information in climate data records (CDRs) is well described in Merchant *et al*., 2017 whose results are based on work done with CCI data.

Type of uncertainty	Acronym
Single sensor uncertainty estimates for every observation	SSEOB
Single sensor uncertainty estimates for CDR	SSECDR
Error covariance matrix for CDR	ERRCOV
L3 merged product	ERRMERG

Table 3. Types of uncertainties for inclusion with CDR datasets, as used by the CCI ECV project teams in CCI Phase 1. The acronyms are used in the tables below.

4.1 Water Vapour

Water vapour (WV) is a component of the hydrological cycle and an important greenhouse gas. The evaporation of water from the Earth's surface is a key process for feedbacks in the climate system such as clouds and radiation, which are a large part of the uncertainty in modelling climate change (IPCC 2013). Water vapour is an important climate variable needed to monitor surface fluxes of moisture, and as such has links with CCI Clouds and CCI Soil Moisture. In particular, there is a need to assess vertical profiles of water vapour in the upper troposphere /

[15] CMUG Phase 2 Deliverable 1.1: User Requirement Document (v0.6), available at http://ensembles-eu.metoffice.com/cmug/CMUG_PHASE_2_D1.1_Requirements_v0.6.pdf.

lower stratosphere where the Outgoing Longwave Radiation (OLR) is most sensitive to changes in WV (Mapes et al, 2017). At the same time, profiles of WV are difficult to measure with current satellite-based passive techniques and new active-based techniques are required to make progress in this area. A product which provides separately UTLS column WV and tropospheric WV could be beneficial.

The user requirements for the satellite-derived Water Vapour ECV as gathered by CMUG are given in Table 4. Eight experts submitted responses on "Water Vapour" to the forty-two questions in the CMUG online survey, and also gave narrative responses providing context to their needs. (The collated written and summarised numerical responses to the CMUG online survey are given in Annex 1.) The data specifications from those answers, plus additional input from experts at ECMWF are given in Table 4, alongside the current GCOS requirements (GCOS 2016). The table lists the desirable data specifications of this ECV for the main applications in climate studies, and the associated acceptable uncertainty in those measurements.

ECMWF is the CMUG partner who takes the lead in working with the Climate Research Group of this ECV project to help research within the CCI to be coordinated, complete and complementary. However, for this ECV there are no modelling or reanalysis studies planned in CMUG Work Packages 3 and 4, instead the CMUG work is in implementing this ECV in the CMF and the ESMValTool (Work Package 5). Thus the modelling and reanalyses requirements stated in Table 4 below will not be validated with CMUG modelling work, but through incorporation in the CMF climate database and the ESMValTool. A version of the Water Vapour CCI dataset will be produced for inclusion in the obs4MIPs archive so that it can be used in the ESMValTool by CMUG as part of its work plan. The current data specification planned by the Water Vapour CCI team meets the obs4MIPs data specification.

Water Vapour as an ECV is not currently available on the Climate Data Store (CDS) of the Copernicus Climate Change Service, or other providers data for climate service. Water Vapour is an important component of the water cycle and will be included in the CDS in future.

The GCOS requirements given in Table 4 are consistent with those gathered by CMUG, although the temporal observing cycle is higher than that required by climate modelling or reanalyses. Thus, if the GCOS requirements are met then so are the CMUG requirements.
In particular, the requirement of an hourly observing cycle would be helpful in quantify the daily cycle of WV.

Parameter	Application	Horizontal Resolution	Observing cycle	Length of record	Error in measured value	Stability of measured values	Error Type (see Table 3)	Source
Water Vapour	GCOS 2016	25 km/NA	Hourly		RH 5%	0.3%/decade	Not stated	GCOS
	Trend monitoring Global/ Regional	10km	1month		RH 1%	0.5%/decade	ERRMERG	ECMWF, CMUG survey

(relative humidity)	Decadal / Seasonal forecasting	25km	1 week / 1 day		RH 1%	0.5%/decade	SSEOB/ ERRMERG	ECMWF, CMUG survey
	Global reanalyses	25 km	1 day		RH 1%	0.5%/decade	SSEOB/ ERRMERG	ECMWF, CMUG survey
	Regional reanalyses	7 km	4 hourly		Dew Point 0.1K	0.5%/decade	SSEOB/ ERRMERG	ECMWF, CMUG survey
Water Vapour (volume mixing ratio)	Climate model evaluation	10 km	1 day / 1 month		10%	0.5%/decade	ERRMERG	ECMWF, CMUG survey
Total column water vapour	GCOS 2016	25 km/NA	4 hourly		2%	0.3%/decade	Not stated	GCOS
	Climate model evaluation	100 km	1 day / 1 month		5%	0.08 kg/m^2 / decade	ERRMERG	ECMWF, BSC, Loew *et al.* (2017)
UTLS column water vapour	Climate model evaluation	100 km	1 day/ 1 month		5%	0.08 kg/m2 / decade	ERRMERG	ECMWF (from values for TCWV)

Table 4. Requirements for satellite derived water vapour observations.

4.2 Sea State

Sea surface state is the fluctuation in the ocean's surface consisting of waves and swell. Monitoring changes in sea state is important because this affects air-sea exchanges of momentum, moisture and CO_2; surface albedo; and the growth/decay of sea ice. These factors impact on marine safety, transport and damage to structures. Sea surface state is included in NWP models, and will be routinely included in climate models in the near future.

The user requirements for the satellite-derived Sea State ECV gathered by CMUG are given in Table 5. Three experts submitted responses on "Sea State" to the forty-two questions in the CMUG online survey, and also gave narrative responses providing context to their needs. (The collated written and summarised numerical responses to the CMUG online survey are given in Annex 1.) The data specifications from those answers, plus additional input from experts at the Met Office are given in Table 5, alongside the current GCOS requirements (GCOS 2016). The table lists the desirable data specifications of this ECV for the main applications in climate studies and reanalyses, and the associated acceptable uncertainty in those measurements.

The Met Office takes the lead in working with the Climate Research Group of this ECV project with the aim that research within the CCI is coordinated, complete and complementary. This ECV was used in one CMUG modelling experiment in Work Package 3 on biophysical feedbacks in the global oceans using the ocean component of the UKESM1. The experiment performed a comparison between model behaviour with level 3 sea state input to the air-sea

CO_2 flux parameterisation, with the model's standard parameterisation which uses just wind speed. The data specification provided in Table 5, for 'Assimilation' is sufficient for the needs of this work.

The CCI Sea State team will be producing a dataset for the obs4MIPs initiative, and the specification for that dataset should be within the data specifications described in Table 5.

Sea State as an ECV is not currently available on the Climate Data Store (CDS) of the Copernicus Climate Change Service. However, it is an important diagnostic of marine / atmosphere kinetic interactions and is provided through CMEMS as wave height at hourly intervals.

The GCOS requirements given in Table 5 are consistent with those gathered by CMUG, although the temporal observing cycle is higher than that required by climate modelling or reanalyses. The spatial resolution of the GCOS requirements is lower than that for some of the applications for modelling or climate monitoring. The error and stability information in the GCOS requirements are also lower than those specified for modelling and monitoring applications.

Parameter	Application	Horizontal Resolution	Observing Cycle	Error in measured value	Stability	Error Type (see Table 3)	Source
Wave height	GCOS 2016	25km	3 hourly	10cm	5cm	Not stated	GCOS
	Trend monitoring Global/Regional	10km	1month	5cm	1cm/ decade	ERRMERG	Met Office,
	Decadal forecasting	25km	1 month	5cm	1cm/ decade	SSEOB/ ERRMERG	Met Office,
	Seasonal forecasting	25 km	1 month	5cm	1cm/ decade	SSEOB/ ERRMERG	Met Office,
	Global reanalyses	25 km	1 day	5cm	1cm/ decade	SSEOB/ ERRMERG	Met Office, CMUG survey
	Regional reanalyses	7 km	1 day	5cm	1cm/ decade	SSEOB/ ERRMERG	Met Office, CMUG survey
	Assimilation	10km	4 hourly	5cm	N/A	SSEOB/ ERRMERG	Met Office,

Table 5. Requirements for satellite derived Sea Surface State observations.

4.3 Above Ground Biomass

Above Ground Biomass ("vegetation") is part of the climate system through its exchange of CO_2, water and energy with the atmosphere, which varies seasonally and regionally. It is also subject to direct intervention by human activity. The data outputs of this CCI+ ECV project are of interest to the budget closure project on the carbon cycle (RECCAP2).

The user requirements for the satellite-derived Above Ground Biomass ECV (AGB) gathered by CMUG are given in Table 6. Two experts submitted responses on "Biomass" to the forty-two questions in the CMUG online survey, and also gave narrative responses providing context to their needs. (The collated written and summarised numerical responses to the CMUG online survey are given in Annex 1.) The data specifications from those answers, plus additional input from experts at the Met Office are given in Table 6, alongside the current GCOS requirements (GCOS 2016). The table lists the desirable data specifications of this ECV for the main applications in climate studies and reanalyses, and the associated acceptable uncertainty in those measurements.

Common requirements for an annual observing cycle, <20% error in the measured value and 10% stability of the data are noted across all the applications specified by users. Requirements for horizontal resolution varied across the different applications, from 500m (GCOS 2016) to 25km (Global reanalyses), although 1km was consistent for all applications except Global reanalyses which noted a requirement for 25km horizontal resolution. 10km horizontal resolution is also a popular requirement for applications involving Trend monitoring (Global/Regional), Decadal forecasting, Seasonal forecasting and Regional reanalyses.

Parameter	Application	Horizontal Resolution	Observing cycle	Error in measured value	Stability	Error Type (see Table 3)	Source
Mass of live and/or dead organic matter in terrestrial vegetation	GCOS 2016	500m - 1km (based on satellite observations of 100-200m)	Annual	<20% error for biomass values >50 t/ha, and <10% for biomass values ≤ 50 t/ha	10%	No agreed standards but see: GOFC-GOLD (2016)	GCOS
	Trend monitoring Global/Regional	1 - 10km	Annual	<20%	10%	ERRMERG	Met Office,
	Decadal forecasting	1 - 10km	Annual	<20%	10%	SSEOB/ ERRMERG	Met Office,
	Seasonal forecasting	1 - 10km	Annual	<20%	10%	SSEOB/ ERRMERG	Met Office,
	Global reanalyses	25km	Annual	<20%	10%	SSEOB/ ERRMERG	Met Office, CMUG survey
	Regional reanalyses	1 - 10km	Annual	<20%	10%	SSEOB/ ERRMERG	Met Office, CMUG survey

Table 6. Requirements for satellite-derived Above Ground Biomass observations.

4.4 Land Surface Temperature

The Land Surface Temperature (LST) is the skin temperature of ground, or of the vegetation canopy if the ground is obscured. It is influenced by the surface albedo, the vegetation cover and the soil moisture. Monitoring land surface temperature is important for understanding land-atmosphere exchange processes and global and regional land surface temperature changes, and for constraining surface energy budgets and model parameters. High-resolution data on LST further allow one to evaluate how realistically climate models represent temperature extremes, which is of use for climate services. One possible application of satellite derived LST data is for quality analysis of in-situ surface temperature measurements, although this is not planned within the current phase of the CCI LST project.

Like other ECVs, LST has a marked diurnal cycle and time information is needed. The target time sampling should be hourly and geostationary satellites have to be used as much as possible. Since most LST products are generated from infrared sensors, only cloud-free observations are available. This makes the comparison with models difficult, especially when cloud-free observations are aggregated at a low spatial resolution (e.g. 25 km) for comparison with LST simulated by a global land surface model (LSM). The latter may implicitly include sub grid cloudy conditions and be lower than the observed LST.

Requirements for an ideal product would be: 1 km, 1 hour, 1 K accuracy, 0.1 K/decade stability. We know this is not feasible in all conditions. According to the Copernicus Global Land Service (CGLS), the best achievable global performance from geostationary satellites is: 5 km, 1 hour, 2 K accuracy, 1 K /decade stability (https://land.copernicus.eu/global/sites/cgls.vito.be/files/products/CGLOPS1_ATBD_LST-V1.2_I1.41.pdf). This means that accuracy and stability are key issues for this ECV.

Parameter	Application	Horizontal Resolution	Observing Cycle	Error in measured value	Stability of measured values	Error Type (see Table 3)	Source
LST (Aggregated radiometric surface temperature of the ensemble of components within the sensor field of view)	GCOS 2016	1km	3 hourly	1K	<0.1K / decade		GCOS
	Trend monitoring Global/Regional	10km/1km	1 week	0.1K	0.01K/ decade	ERRMERG	Met Office,
	Decadal forecasting	25km	1 week	0.1K	0.1K/ decade	ERRCOV	Met Office,
	Seasonal forecasting	25 km	12 h	0.1K	0.05K/ decade	ERRCOV	Met Office,
	Global reanalyses	25 km	12 h	0.1K	0.01K/ decade	ERRCOV	Met Office, CMUG survey
	Regional reanalyses	5 km	6 h	0.1K	0.01K/ decade	ERRCOV	CMUG survey
	Assimilation	5 km	6 h	0.1K	0.01K/ decade	ERRCOV	CMUG survey

Table 7. Requirements for satellite derived Land Surface Temperature observations.

4.5 Permafrost

This ECV includes the temperature distribution in the permafrost layer and the depth of the overlying active layer where seasonal freezing and thawing occur. Frozen ground at high altitudes and high latitudes reacts sensitively to climate and environmental changes, particularly changes in air temperature and snow. Changes in the seasonal frozen ground can impact the physical state of the permafrost including terrain stability and hydrology. These will have subsequent impacts on the carbon cycle. Development of specific understanding of these processes and their interactions requires a data product at temporal resolutions of daily to weekly and very high spatial resolutions of the order of tens of meters for a range of bio-climate zones. However, for applications such as climate model development and evaluation, both the spatial and temporal requirements are less demanding, in part because the climate models are run at much lower resolutions and still have significant biases (Burke *et al*., 2020). Permafrost early warning systems which forecast the state of the permafrost at seasonal to decadal time scales can be improved through the inclusion of permafrost temperatures. This will require measurements of higher accuracy than for climate model evaluation (~0.1 K).

Permafrost physics directly impacts the carbon cycle. Process-based studies examining the interaction between the soil and vegetation biogeochemistry and physical state of the permafrost require high spatial and temporal resolution. The smaller the error in the physical state of the permafrost the easier it is to understand the contribution of uncertainties in the estimation of the carbon cycle. There is also a requirement to quantify global carbon budgets for both the present and future projections (Gasser *et al*., 2018). The relative contributions of methane and carbon dioxide from permafrost regions to these budgets needs to be quantified (Turetsky *et al*., 2020). Data outputs of this CCI+ ECV project are of interest to the budget closure project on the carbon cycle (RECCAP2). A spatial resolution ~25 km would be appropriate when quantifying global carbon budgets.

Some of the applications shown in Table 8 are relatively novel, particularly the seasonal/decadal forecasting and detection and attribution components. Therefore precision/accuracy/stability requirements are an estimate which will be refined further once these applications are more mature.

Parameter	Application	Horizontal Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Thermal state of permafrost	GCOS 2016	Sufficient sites to characterise each bio-climate zone	Daily to weekly	0.1K	0.1K	0.01K/ decade	ERRMERG	GCOS 2016
	Climate model development and evaluation	25km	1 month to 1 year	0.3K	0.3K	0.5K/ decade	ERRMERG	Met Office
	Trend monitoring Global/Regional	10km/1km	1 week to 1 month	0.1K	0.1K	0.01K/ decade	ERRMERG	
	Seasonal / Decadal forecasting	25km	1 week to 1 month	0.1K	0.1K	0.1K/ decade	ERRMERG	
	Reanalyses	25 km	12 h	0.1K	0.1K	0.05K/ decade	ERRMERG	
	Long term trend detection and attribution	25km	1 month to 1 year	0.2K	0.2K	0.1K/ decade	ERRMERG	Met Office
	Carbon budgets	1km-25km	1 week to 1 month	0.1K	0.1K	0.1K/ decade	ERRMERG	Met Office
	Climate-carbon process studies	1km-25km	Daily to 1 week	0.1K	0.1K	0.1K/ decade	ERRMERG	Met Office

Active Layer Thickness	GCOS 2016	Sufficient sites to characterise each bio-climate zone	Daily to weekly	2cm	2cm			GCOS 2016
	Climate model development and evaluation	25km	1 month to 1 year	10cm	10cm	10 cm/ decade	ERRMERG	Met Office
	Long term trend detection and attribution	25km	1 month to 1 year	1cm	1cm	1cm/ decade	ERRMERG	Met Office
	Carbon budgets	1km-25km	1 week to 1 month	1cm	1cm	1cm/ decade	ERRMERG	Met Office
	Climate-carbon process studies	1km-25km	Daily to 1 week	1cm	1cm	1cm/ decade	ERRMERG	Met Office

Table 8. Requirements for satellite derived Permafrost observations.

4.6 Lakes

Lakes and inland seas play an important role in global and local climate regulation. As they are distributed worldwide, they will be impacted in different ways by climate change. They are also an essential resource, providing livelihoods for communities around the globe.

Reanalysis of lake parameters (ECVs) could be very useful in the context of regional climate modelling, in particular to refine the setup of evaluation runs, whose purpose is the assessment of Regional Climate Model (RCM) skill using quasi-observed atmospheric boundary conditions, and prognostic sea surface temperature and sea ice fraction (Giorgi & Gutowski, 2016). The inclusion of prescribed lake variables would work as an observational constraint, particularly useful for RCMs which rely on their land surface schemes for a basic description of lakes, but also very useful for RCMs with more refined lake components. In this context, modellers are expecting data at high temporal frequency (at least daily) for the assessment of lake components in RCM, with the additional requirement of continuity (i.e. no gaps) for their use to prescribe lake properties in RCM simulations. For this purpose, lake reanalysis should also be available on sufficiently long periods to allow climatological analysis, with a good overlap with existing reanalyses, and consistent with them in terms of resolution and accuracy, including the accurate representation of existing trends on lake temperature and ice fraction.

The main parameters required for shorter timescale NWP modelling (seasonal/decadal) are lake extent and depth for ancillary input. In future, depth and extent changes may also be required although current models (e.g. Met-UM) can't handle these at present. Surface temperature is the main output of the FLake module used in the Met-UM, so observations of this for validation as well as input would be useful along with ice cover and thickness. Detailed information would be required on how the parameters were obtained. If colour could be related to opacity/extinction coefficient then it could be useful also. Mixed-layer thickness is much harder to obtain even than ice thickness but might also be interesting if possible. In the future it will also be relevant to distinguish between lakes which are fresh water and those which are salty water. Lake colour will be useful when carbon budget modelling starts to take account of the effect of lakes, it may also be useful for modelling stratification or mixing below the lake surface in future.

A Global Lakes Data Base (GLDB) has been set up for NWP applications containing information about lake location (latitude, longitude), water surface area, and lake mean and max depths. The mean depth is provided as a gridded data set (Toptunova, *et al.*, 2019). This is the sort of information which is most useful on NWP scales.

Given evidence that forecast skill is improved by initialising land surface parameters, soil moisture in particular, for seasonal (Seo *et al.*, 2019) and sub-seasonal (Seneviratne *et al.*, 2012) forecasts it seems likely that lakes parameters could improve forecasts in a similar way. The large thermal capacity of these systems could provide mechanisms for long term memory in the climate system also providing crucial information for hindcasts and reanalyses.

Parameter	Application	Horizontal Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see table 3)	Source
Lake water level	GCOS 2016	100 m	Daily	3cm for large lakes; 10cm for the remainder	3cm for large lakes; 10cm for the remainder	1cm / decade		WMO
	Trend monitoring	1km / 100m	Daily	3cm for large lakes; 10cm for the remainder	3cm for large lakes; 10cm for the remainder	1cm / decade	ERRMERG	MO
	Seasonal / Decadal forecasting	25 km	1 week	3cm for large lakes; 10cm for the remainder	3cm for large lakes; 10cm for the remainder	1cm / decade	SSEOB	MO
	Reanalyses	10 km	12 h	3cm for large lakes; 10cm for the remainder	3cm for large lakes; 10cm for the remainder	1cm / decade	ERRCOV	MO
Water extent	GCOS 2016	20m	Daily	10% (relative); 5% (for 70 largest lakes)	10% (relative); 5% (for 70 largest lakes)	5% / % / decade		WMO
	Trend monitoring	100 m	Daily	10%	10%	5% / decade	ERRMERG	MO
	Seasonal / Decadal forecasting	10 / 25 km	1 week	10%	10%	5% / decade	SSEOB	MO
	Reanalyses	10 / 25 km	Daily	10%	10%	5% / decade	ERRCOV	MO
Lake surface water temperature	GCOS 2016	300m	Weekly	1K	1K	0.1K/decade		WMO
	Trend monitoring	300 m	Daily	0.1K	0.1K	0.01K/ decade	ERRMERG	CMUG survey
	Seasonal / Decadal forecasting	10 / 25 km	1 week	0.1K	0.1K	0.05K/ decade	SSEOB	CMUG survey
	Reanalyses	10 / 25 km	Daily	0.1K	0.1K	0.05K/ decade	ERRCOV	CMUG survey
Lake ice thickness	GCOS 2016	100m	Monthly	1-2cm	1-2cm			WMO
	Trend monitoring	<200 m	Weekly	5cm	5cm		ERRMERG	MO
	Seasonal / Decadal forecasting	10 / 25 km	1 week	10cm	10cm		SSEOB	MO

	Reanalyses	10 / 25 km	Daily	10cm	10cm		ERRCOV	MO
Lake Ice Cover	GCOS 2016	300m	Daily	10%	10%	1%/decade		WMO
	Trend monitoring	200 m	Weekly	10%	10%	1%/decade	ERRMERG	MO
	Seasonal / Decadal forecasting	10 / 25 km	Daily	10%	10%	1%/decade	ERRCOV/SSEOB	MO
	Reanalyses	10 / 25 km	Daily	10%	10%	1%/decade	ERRCOV	MO
Lake Colour	GCOS 2016	300m	Weekly	30%	30%	1%/decade		WMO
	Trend monitoring	300 m	Weekly	30%	30%	1%/decade	ERRMERG	MO
	Seasonal / Decadal forecasting	10 / 25 km	Weekly	30%	30%	1%/decade	ERRCOV/SSEOB	MO
	Reanalyses	10 / 25 km	Daily	30%	30%	1%/decade	ERRCOV	MO

Table 9. Requirements for satellite derived Lake observations.

4.7 Snow

Snow cover plays a significant role in the climate system, covering 50% of the northern hemisphere’s land surface in the winter. It influences the climate system through feedbacks such as high albedo and heat insulation, and contributes to soil moisture and runoff. It is also a vital freshwater resource, and a seasonal component in the water cycle especially in high- and mid-latitude regions.

Requirements for an ideal snow cover fraction product would be: 1 km, daily, 5 % accuracy. A vegetation correction over forested areas is needed. Also a correction for complex topography within a pixel. We will focus on snow cover fraction in a first stage. The non-GCOS lines in the Table below account for the information collected in the framework of the CMUG user requirement survey (9 users addressed snow variables).

Parameter	Application	Horizontal Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Area covered by snow	GCOS 2016	1km (100m in complex terrain)	Daily	5% (maximum error of omission and commission in snow area). Location accuracy better than 1/3 IFOV with target IFOV 100 m in areas of complex terrain, 1 km elsewhere	5% (maximum error of omission and commission in snow area). Location accuracy better than 1/3 IFOV with target IFOV 100 m in areas of complex terrain, 1 km elsewhere	4% (maximum error of omission and commission in snow area); location accuracy better than 1/3 IFOV with target IFOV 100 m in areas of complex terrain, 1 km elsewhere		WMO (2008c) IGOS (2007), IACS/ UNESCO, 2009
	Trend monitoring	Goal 1km, Threshold 25km	Daily	Goal 5% Threshold 10% (of std)	Goal 5% Threshold 10% (of std)	Goal 1%/decade Threshold 2% / decade (of std)	SSECDR ERMERG	CMUG user survey
	Seasonal / Decadal forecasting	Goal 1km, Threshold 25km	Daily	Goal 5% Threshold 10% (of std)	Goal 5% Threshold 10% (of std)	Goal 1%/decade Threshold 2% / decade (of std)	SSECDR ERMERG	CMUG user survey

	Reanalyses	Goal 1km, Threshold 25km	Daily	Goal 5% Threshold 10% (of std)	Goal 5% Threshold 10% (of std)	Goal 1%/decade Threshold 2% / decade (of std)	SSECDR ERMERG	CMUG user survey
Snow depth	GCOS 2016	1km (100m in complex terrain)	Daily	10mm	10mm	10mm		WMO (2008c) IGOS (2007), IACS/ UNESCO, 2009
	Trend monitoring	Goal 1km, Threshold 25km	Daily	Goal 5% Threshold 10% (of std)	Goal 5% Threshold 10% (of std)	Goal 1% /decade Threshold 2% /decade (of std)	SSECDR ERMERG	CMUG user survey
	Seasonal / Decadal forecasting	Goal 1km, Threshold 25km	Daily	Goal 5% Threshold 10% (of std)	Goal 5% Threshold 10% (of std)	Goal 1%/decade Threshold 2% / decade (of std)	SSECDR ERMERG	CMUG user survey
	Reanalyses	Goal 1km, Threshold 25km	Daily	Goal 5% Threshold 10% (of std)	Goal 5% Threshold 10% (of std)	Goal 1%/decade Threshold 2% / decade (of std)	SSECDR ERMERG	CMUG user survey
Snow water equivalent	GCOS 2016	1km	Daily	10mm	10mm			WMO (2008c) IGOS (2007), IACS/ UNESCO, 2009
	Trend monitoring	Goal 1km, Threshold 25km	Daily	Goal 5% Threshold 10% (of std)	Goal 5% Threshold 10% (of std)	Goal 1%/decade Threshold 2% / decade (of std)	SSECDR ERMERG	CMUG user survey
	Seasonal / Decadal forecasting	Goal 1km, Threshold 25km	Daily	Goal 5% Threshold 10% (of std)	Goal 5% Threshold 10% (of std)	Goal 1% /decade Threshold 2% / decade (of std)	SSECDR ERMERG	CMUG user survey
	Reanalyses	Goal 1km, Threshold 25km	Daily	Goal 5% Threshold 10% (of std)	Goal 5% Threshold 10% (of std)	Goal 1%/decade Threshold 2% / decade (of std)	SSECDR ERMERG	CMUG user survey

Table 10. Requirements for satellite derived Snow observations.

4.8 Sea Surface Salinity

Sea surface salinity (SSS) is an indicator of the balance of the global water cycle e.g. evaporation, precipitation, glacier and river runoff. SSS plays a key role in determining the large-scale density gradients alongside sea surface temperature, provides an expression of ocean frontal features and eddies and influences the deep-water formation processes that control the global ocean circulation. Hence it is important for climate modellers interested in variability and prediction at decadal and longer timescales. This includes climate projection studies in which the future role of a potential sea water freshening on the Atlantic circulation is still unclear. In this context, the development and temporal extension of new improved SSS observations in high latitudes, and in particular in the regions where deep-water convection happens (e.g. the Labrador and Nordic Seas in the Northern Hemisphere or the Weddell and Ross Seas in the Southern Hemisphere) would be useful for several purposes, from climate model evaluation, to model tuning, or the generation of in-house ocean reanalyses for forecast initialization. High resolution SSS observations can be particularly leveraged to evaluate model output.

On large scales, sea surface salinity can also be used to infer long-term changes of the global hydrological cycle. Knowledge of SSS helps to understand the evaporation and precipitation estimates for the global ocean; river runoff and glacial and land ice melting rates; sea surface density, alkalinity and dissolved CO_2, and additionally allows to track eddy activity and diagnose eddy contributions to the transport of salt in high-resolution model configurations to evaluate their realisms. SSS is itself influenced by other ocean, atmospheric and ice variables, such as sea surface temperature, sea surface winds, precipitation, evaporation, ice-sheet melting and sea ice concentrations. Salinity of seawater is a dimensionless measurement and is expressed in practical salinity units, or psu (PSS-78, Lewis and Perkin, 1978)

Parameter	Application	Horizontal Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Sea surface salinity (Bulk surface salinity, skin surface salinity, near surface salinity at stated depth)	GCOS 2016	1-100 km	Hourly to monthly	0.01 psu	0.01 psu	0.001 psu	See EOV specification sheet at www.goosocean.org/eov	CMUG survey
	Trend monitoring Global/Regional	100km/ 10km	1 week	0.01 psu	0.01 psu	0.001 psu	ERRMERG	CMUG survey
	Decadal forecasting	25km	1 week	0.01 psu	0.01 psu	0.001 psu	SSEOB / ERRMERG	BSC
	Seasonal forecasting	25 km	daily	0.01 psu	0.01 psu	0.001 psu	SSEOB / ERRMERG	BSC
	Global reanalyses	25 km	daily	0.01 psu	0.01 psu	0.001 psu	SSEOB / ERRMERG	BSC
	Regional reanalyses	5 km	daily	0.01 psu	0.01 psu	0.001 psu	SSEOB / ERRMERG	Met Office
	Assimilation	5 km	daily	0.01 psu	0.01 psu	0.001 psu	SSEOB / ERRMERG	BSC
	High resolution Model Evaluation	5km	1 week	0.01 psu	0.01 psu	0.01 psu	SSEOB / ERRMERG , BSC	

Table 11. Requirements for satellite derived Sea Surface Salinity observations.

4.9 Land Cover

Land cover (LC) is the bio-physical cover on the Earth's surface. It describes the spatial distribution of vegetation (forests, grasslands, croplands, etc.) and man-made features (living space, agriculture and forestry). LC has a key role in the energy, water and carbon fluxes exchanged between the land surface and the atmosphere through distribution and changes of vegetation and soil carbon, or biophysical effects such as sunlight reflection and evapotranspiration, which, in turn, affect surface temperature and precipitation. Furthermore, LC plays a pivotal part also in the characterization of sources and sinks of gases and aerosols, since land acts both as a source and a sink of greenhouse gases, and mineral dust emissions or wildfires, for example, are strongly linked to changes in land cover. Hence LC changes can have a significant impact on the environment and climate at many different spatial scales (locally, regionally and globally).

Detailed information about land cover is therefore crucial for global and regional climate modelling over a range of timescales. Earth system models are the most advanced tools with which to conduct studies on climate monitoring/attribution, and also to predict future climate. Land cover information is used in climate models for initialization as well as to prescribe boundary conditions. The current generation of earth system models includes different broad categories of vegetation, referred to as plant functional types (PFTs), sharing similar characteristics and functional role in the ecosystem. A conversion ("cross-walking") scheme translates the categorical LC classes to their PFT fraction equivalent. The land cover information is then translated into surface parameters (e.g. albedo, LAI, surface roughness, fractional vegetation cover), which provide the lower boundary condition for the atmospheric models. In addition to full Earth system simulations, detailed regional land cover information provides very valuable information for process studies such as the assessment of the impact of fires. LC information is also widely used to help model development and validation.

In the context of the CCI+ phase of this project, a review has been conducted to survey users of LC data. Such data are used in either land-only models, atmospheric models or full Earth System models, for purposes that range from climate modelling, studies and services, climate and environment monitoring and attribution to NWP and atmospheric reanalyses. Some of the key processes of interest for those users are water cycle, snow and ice energy budgets, drought, dynamic vegetation, albedo, land-surface-atmosphere interaction. The requirements expressed by users confirm the need for an increase in resolution.

Climate modellers who use LC data have indicated that while the medium resolution (MR) LC is better able to identify vegetation types (e.g., needleleaf versus broadleaf, or tree versus shrub) due to a better temporal resolution (more frequent satellite repeat cycle), products from the High Resolution Land Cover (HRLC) CCI ECV project are of interest if they help with understanding the climate system processes represented in models, such as the fractional coverage of Plant Functional Types (PFTs) within a MR grid cell (typically a 300m cell). High resolution information could be used to calibrate cross-walking tables used to convert land cover into land surface tile fractions for use in models, and which can have quite a range of uncertainty in the legend definition. Hence, the combination of the MR and HR LC is potentially very useful.

Furthermore, high resolution is also needed in order to better assess and attribute land cover changes within a MR pixel and to capture land use changes that are missed by MR observations. The observing cycle for the LC ECV is required to be shorter than the 2-5 years previously specified. This is most relevant for land cover change detection as those changes occur in timeframes shorter than 2 years.

The data outputs of this CCI+ ECV project are of interest to the budget closure project on the carbon cycle (RECCAP2).

The requirements for satellite derived observations of land cover are given in Table 12.

Parameter	Application	Horizontal Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source

(HR) Land cover type	model development, initialization, reanalysis, verification, surface energy budget	(10-100m) 100-1000m	Daily Monthly Yearly	< 10%	< 10%	Consistency should be maintained across several consecutive maps	ERRMERG	Met Office, CMUG survey
Land cover change	trend monitoring	(10-100m) 100-1000m	Monthly Yearly <2 yr	< 10%	< 10%	Consistency should be maintained across several consecutive maps	ERRMERG	Met Office, CMUG survey

Table 12. Requirements for satellite derived High Resolution and Medium Resolution Land Cover observations.

4.10 Ice Sheets

The Greenland and Antarctic ice sheets are separate ECVs but are considered together in this document because requirements are identical.

The Greenland Ice Sheet has shown rapid change in the last decade, indicated by thinning along the margins, accelerating outlet glaciers and increasing mass loss (http://esa-icesheets-greenland-cci.org/). Changes in the Antarctic ice shelf have previously contributed to changes in climate and sea level rise.

The ice sheets are interesting to climate modellers because they interact with other components of the climate system (e.g. freshwater fluxes from ice sheets affecting sea-level and ocean circulation, orographic forcing of winds, and albedo). Ice sheet data (from satellites or elsewhere) are exploited in global climate models for ice sheet initialisation and recently to constrain uncertainties in probabilistic ice sheet projections. They are used in regional models that specialise in ice interactions in the climate system such as RACMO[16], HIRHAM5 and MAR. The requirements from both regional and global modellers are captured in Table 13. Not all relevant parameters have quantified requirements (and are hence not included in Table 13). The absolute ice thickness of ice sheets has a significant impact on the initialisation, the bedrock topography underneath the ice sheets has also been shown to significantly influence the ice sheet behaviour. The potential for glacial isostatic adjustment and geothermal heat flux are very challenging to measure but should nevertheless be mentioned as scientifically interesting. Surface melt of the ice sheets (as part of the surface mass balance) and ice shelves (as potential pre-processor of hydrofracturing) is an important variable for trend monitoring and initialisation. The Antarctic surface accumulation rate is discussed as potential cause of Antarctic mass gain in a warmer climate, until increased surface melt dominates the response. Monitoring trends in Antarctic wide surface accumulation and melt would allow to constrain short term (decadal to century scale) prediction of the Antarctic ice sheet.

[16] https://www.projects.science.uu.nl/iceclimate/models/greenland.php

The precision column has not been populated based on feedback from users that without "truth" for assessment this metric was not useful. This feedback also reflected a dissatisfaction with the assumption that gravimetric satellites provide a best measurement, stating that differences in processing technique can provide different answers. It was felt that the measurement error is always underestimated and if the true uncertainty were used the accuracy and stability requirements could not be met.

Parameter	Application	Horizontal Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Surface elevation change	GCOS 2016	Horizontal 100m	30 days	N/A	0.1m/year	0.1m/year		GCOS
	Trend monitoring	<500 m	monthly	N/A	<0.1m/yr	<0.1 m/ decade	SSEOB	Met Office
	Initialisation	<5km	annual	N/A	0.1m/yr	<0.1m/yr	SSEOB	Met Office
Ice Velocity	GCOS 2016	Horizontal 100m	30 days	N/A	0.1m/year	0.1m/year		GCOS
	Trend monitoring	0.1m/yr	monthly	N/A	<10m/yr	<10m/yr	SSEOB	Met Office
	Initialisation	0.5m/yr	annual	N/A	30 m/yr	<10m/yr	SSEOB	Met Office
Calving Front Location	Initialisation	<5km	annual	N/A	0.1m/yr	<10m/yr	SSEOB	Met Office
Grounding Line Location and Thickness	GCOS 2016	Horizontal 100m; vertical 10m	annual	N/A	1m	10m		GCOS
	Initialisation	<5km	annual	N/A	1m	10m	ERRCOV	Met Office
Gravimetric Mass Balance	Trend monitoring	50 km	30 days	N/A	$10km^3$/year	$10km^3$/year	ERRMERG	Met Office

Table 13. Requirements for satellite derived Ice Sheet observations.

4.12 Sea Surface Temperature

Sea surface temperature (SST) is an important variable to monitor over many timescales as a key indicator of climate change. Satellite SST data are crucial to obtaining globally complete SST analyses and in particular the high temporal and spatial resolution that is increasingly needed for understanding processes such as ENSO, NAO, PDO etc.

The IPCC AR5 report states *"Since the AR4, major improvements in availability of metadata and data completeness have been made, and a number of new global SST records have been produced. Intercomparisons of new SST data records obtained by different measurement methods, including satellite data, have resulted in better understanding of uncertainties and biases in the records."* and so removal of the biases and understanding biases is clearly a critical need for climate monitoring. It is also important for climate change to monitor the SSTs over the Arctic Ocean as there is a lack of conventional air temperature measurements in the Arctic.

To be able to use an SST data set as a boundary condition for atmospheric reanalyses or in atmosphere-only climate simulations, gridded data sets with complete coverage over the global ocean are needed. These are based on a special form of Optimal Interpolation that retains large-scale correlation structures and can accommodate sparse data coverage. The OSTIA SST analysis is used by the Met Office and other NWP centres for both operational forecasting (NWP and Ocean) and an OSTIA reanalysis has been run using the historical observations available. This complements the HadISST climate quality data analysis produced in the MOHC which makes use of the CCI SST climate data records. These high resolution analyses are linked to the longer-term climate record of SST. HadISST (v2.2.0.0) analysis is being used in ERA5 (along with OSTIA for the more recent period) which uses satellite data (AVHRR and ATSR) from 1979 onwards along with in-situ data.

The requirements for satellite SST are given in Table 15 for a number of applications related to climate modelling. As models improve and are able to resolve finer detail the horizontal and temporal resolution requirements for the data will become ever more stringent. An important consideration is whether sea surface skin temperature or sea surface subskin temperature (also known as a foundation temperature) is required. The requirements are the same for both. For long term trend monitoring both parameters are of interest with foundation temperature used more in the past but for the satellite era skin temperature could also be used and models are being developed to use skin temperature or even radiances. A recent study has demonstrated that using foundation rather than skin temperature in estimates of air-sea CO2 fluxes can result in a significant underestimate (Watson *et al.*, 2020) Long term trend monitoring and attribution is the most challenging application with high demands on the accuracy and stability of the product especially if regional trends are required.

There are a number of requirements for initialising the state of seasonal, decadal and coupled climate model runs which all have similar requirements on accuracy. SST is of particular importance, because (1) it is at the interface with the atmosphere and thus plays a key role enabling the teleconnection mechanisms that provide predictive skill over the continents, and (2) it is more reliably and systematically monitored, both in terms of spatial coverage and temporal frequency. The deep ocean temperatures, for which observations are sparser and have higher uncertainty, are also of great importance, in particular to represent the oceanic processes that provide skill at the longer-range forecasts (through their influence on the surface). For reanalysis the requirement is to provide at least a 3-hourly update to the SST field as a boundary condition for the assimilation of the atmospheric and other oceanic variables.

SST is also an important dataset for Climate Services with many applications (fisheries, military, tourism, transport, etc.). Here a range of horizontal and temporal sampling options will be required for delivery to the diverse list of users. There is a requirement to have the reprocessed SST data within a month of real time to be able to put severe weather events into context for government or media requests.

Application	Horizontal resolution	Temporal sampling	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Trend monitoring (global/regional)	10km/1km	1 week	0.1K	0.1K	0.01K/ decade	ERRMERG	BSC, CMUG survey
Seasonal f/c	20km	12h	0.1K	0.1K	0.05K/ decade	ERRCOV	BSC, CMUG survey

Decadal f/c	20km	1 month	0.1K	0.1K	0.1K/ decade	ERRCOV	BSC, CMUG survey
Climate quality analysis	20km	12h	0.1K	0.1K	0.05K/ decade	ERRCOV	BSC, CMUG survey
Global Reanalysis	20km	12h	0.1K	0.1K	0.01K/ decade	ERRCOV	BSC, CMUG survey
Regional Reanalysis	5km	3h	0.1K	0.1K	0.01K/ decade	ERRCOV	BSC, CMUG survey
Climate services	1km	3h	0.1K	0.1K	0.01K/ decade	ERRMERG	BSC, CMUG survey

Table 14. Requirements for satellite derived *SST observations. The accuracy and stability values assume global coverage for 100km spatial scales.*

The stated requirements are generally the same as the previous set of CMUG user requirements, except that the spatial resolution required for seasonal-to-decadal forecasting and global reanalysis has increased from 30km to 20km. This reflects the increasingly common use of eddy-permitting rather than eddy-parameterising ocean models for these applications. Eddy-resolving models are now being used for operational ocean forecasting, and as available computing resources increase these models will start to be used for reanalysis and seasonal-to-decadal forecasting too. It is therefore likely that in the medium-term the spatial resolution required from SST products for these applications will increase to 10km or finer.

4.13 Ocean Colour

The impact of climate change on marine ecosystems and the ocean carbon cycle, from global to regional scales, can only be quantified by using long-term data sets, including satellite ocean colour. Synoptic fields of ocean colour (derived chlorophyll pigment), are used as an index for phytoplankton biomass, which is the single most important property of the marine ecosystem. Ocean colour is also the basis to infer primary production (CO_2 uptake by algae) and is currently the only source of observational data offering complete global coverage. This offers a wide scope of ocean colour CDRs applications, which include:

- initialisation and verification of coupled ocean-biogeochemical models and potentially ocean-atmosphere-biogeochemical models.
- data assimilation for state, as well as parameter estimation in ocean forecasting models and reanalyses.
- physical-biogeochemical process studies, such as primary production, respiration and interactions at the air-sea interface
- climate process studies, for instance examining the biological carbon pump, phytoplankton phenology, and mechanisms controlling seasonal phytoplankton blooms. Numerous applications for Climate and Marine Services, such as fish stock assessments, carbon sequestration, eutrophication, ecosystem health monitoring and integrated ecosystem assessment to name a few
- trend analysis to assess the impact of climate change and variability on the marine ecosystem
- model boundary conditions, and calculation of ocean albedo.

The CMC requirements for satellite ocean colour observations are given in Table 15. Compared to the GCOS requirements these are close to the goals of GCOS in terms of resolution and observing cycle. The requirements could also be sub-divided into ocean and coastal waters where the first is the easiest case to achieve the stated requirements. Higher spatial and temporal resolution is typically required in coastal waters, at least daily and for some use cases sub-kilometre, while accuracy requirements are harder to achieve due to the optical complexity of Case 2 waters.

The CMC requirements remain largely unchanged from the previous survey carried out by CMUG. The only changes are that the horizontal resolution required for decadal forecasting, seasonal forecasting, and global reanalyses has reduced from 25 km to 20 km. The requirement appears to be for the products to match the resolution of a 1/4° model grid, which translates to 25 km at low-latitudes and 20 km at mid-latitudes. This therefore does not seem to represent any fundamental change in requirements.

Most members of the CMC are content with merged multi-sensor products at daily or lower resolution, which matches the GCOS requirements and current CCI products. For uses such as climate model validation and monitoring, this is sufficient. One respondent to the CMUG survey expressed a preference though for single-sensor products at sub-daily resolution. Their use case was for data assimilation into reanalyses, with single-sensor along-track products providing benefit for high-resolution assimilation techniques.

There are a range of other possible products which could be considered for example in carbon budget assessments, such as phytoplankton functional types and particulate organic carbon. To date, modellers have not expressed any firm requirements for these, but demand may increase in future, once modellers can be confident in the accuracy of such products.

Parameter	Application	Horizontal Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Derived chlorophyll *a*	Trend monitoring Global/Regional	4km	1month	30% or under	30% or under	2%/decade	ERRMERG	CMUG survey
	Decadal forecasting	20 km	1 month	30% or under	30% or under	2%/decade	SSEOB/ ERRMERG	CMUG survey
	Seasonal forecasting	20 km	1 month	30% or under	30% or under	2%/decade	SSEOB/ ERRMERG	CMUG survey
	Global reanalyses	20 km	1 day	30% or under	30% or under	2%/decade	SSEOB/ ERRMERG	CMUG survey
	Regional reanalyses	7 km	1 day	30% or under	30% or under	2%/decade	SSEOB/ ERRMERG	CMUG survey
	Shelf (tidal) seas	4 km to 200 m	1 day	30% or under	30% or under	2%/decade	SSEOB/ ERRMERG	CMUG survey
	Assimilation	4km	1 day	30% or under	30% or under	N/A	SSEOB/ ERRMERG	CMUG survey

Table 15. Requirements for satellite derived ocean colour observations

4.14 Sea Level

Sea level rise through thermal expansion of the oceans is one of the most important consequences of a warming climate and its potential impacts justify a careful study of sea level variability and trends at global, regional and local scales. It is also a key parameter to monitor some important features of climate variability such as the ENSO.

For the CMC a first interest is to run historical realisations of the climate and to compare the modelled regional variability of sea level with that observed, and to inform detection and attribution studies of the observed changes. Getting models to match the observed variability improves confidence in their ability to provide useful predictions. It is also important to evaluate the overall sea level rise due to rising temperatures and melting of ice sheets simulated by models against the observations in order to facilitate weighting of model ensembles and/or the application of observational constraints on model projections of future climate change.

Another interest of the CMC for sea level data concerns data assimilation in ocean models. These data provide invaluable information to complement in-situ observation in order to constrain the simulated ocean circulation. Ocean data assimilation can either be used to initialize the ocean component of the coupled models used for climate prediction or through the use of ocean-reanalyses for a wide range of applications (for example to force atmospheric stand-alone models, to evaluate ocean models, and to analyse climate variability). As the spatial resolution of these models increases higher resolution datasets are required, hence the change in horizontal resolution requirement since the last version of this document.

Sea-level trend analysis and detection/attribution studies also require sea-level ECVs with specific requirements in particular for stability. The data outputs of this CCI+ ECV project are of interest to the Sea Level Budget Closure project. The CRC requirements for satellite sea level observations are given in Table 16.

Parameter	***Application***	***Horizontal Resolution***	***Observing Cycle***	***Precision***	***Accuracy***	***Stability***	**Error Type (see Table 3)**	**Source**
Ocean dynamic topography	*Model Development and Evaluation*	*20 km*	*5 days*	*1 cm*	*1 cm*	*2mm/ decade*	*SSEOB*	*Survey*
	Reanalyses and data assimilation	*20 km*	*2 days*	*1 cm*	*1 cm*	*2mm/ decade*	*SSEOB*	*Survey*
	Long Term Trend Monitoring and Attribution	*20 km*	*2 days*	*1 cm*	*1 cm*	*2mm/ decade*	*SSEOB*	*Survey*
Coastal sea level change	*Model Development and Evaluation*	*10 km*	*5 days*	*1 cm*	*1 cm*	*2mm/ decade*	*SSEOB*	*Survey*
	Reanalyses and data assimilation	*10 km*	*2 days*	*1 cm*	*1 cm*	*2mm/ decade*	*SSEOB*	*Survey*
	Long Term Trend Monitoring and Attribution	*10 km*	*2 days*	*1 cm*	*1 cm*	*2mm/ decade*	*SSEOB*	*Survey*

Table 16. Requirements for satellite derived sea level observation.

4.15 Sea Ice

Sea Ice is a central element of the Polar climate, amplifying the local warming due to climate change through its key role in the (positive) ice albedo feedback. The recent years have seen a strong decrease in the Arctic sea-ice cover while the Antarctic sea-ice extent remained stable with no statistically significant trends (IPCC, 2019), raising scientific interest to better understand the reasons for the different ongoing changes. For this purpose, reliable observational data of the sea-ice properties covering both polar regions is crucial. Such data is only available from satellites, which are the main source of information used by scientists to study the large-scale evolution and predictability of sea-ice.

The most important sea-ice properties that can be obtained from satellites are sea-ice concentration (derived from passive-microwave retrievals from 1979 onwards) and sea-ice thickness (derived more recently from laser or radar altimetry and, for thin ice, from SMOS). Additionally, but much less widespread, sea ice drift speed and albedo products are also developed and can be used to understand the dynamics of sea ice, and also for process-based studies focused on the importance of sea ice for Polar Amplification. From ice concentrations, integrative quantities such as sea-ice area or sea-ice extent can be derived, while sea-ice volume can be obtained combining sea-ice thickness and sea-ice concentration data. The main specifications for data parameters under different applications are shown in Table 17.

Three important limitations for the usage of sea-ice data for climate model applications (such as model-evaluation and model-initialization) are the short length of certain records (e.g. the sea ice thickness products), the poor description of uncertainties and uncertainty correlations for both sea ice thickness and concentration retrievals, and in some cases the poor consistency between records based on different sensors, as they all hinder our knowledge of the long-term and interannual sea ice evolution.

Most climate modellers use level 3 data, with level 1 or level 2 data being primarily used for algorithm development. This focus on level 3 data might, however, change with the ongoing development of satellite simulators of sea ice that aim at directly providing level 2 fields from the model simulations.

The merging of several products into a single field is not strictly necessary but might be useful for sea-ice thickness where Cryosat provides information on thicker ice and SMOS provides information on thin ice. In any case, it should always be possible to trace back the underlying data source at each grid point.

Parameter	Application	Horizontal Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Sea-ice cover (first year & multi-year ice)	trend monitoring Global/Regional	10km / 10km	1 day	5%	5%	1%/decade	SSEOB	BSC, CMUG survey
	decadal f/c	50km	1 month	5%	5%	1%/decade	SSEOB	BSC, CMUG survey
	Initialise	5km	1 day	5%	5%	1%/decade	SSEOB	BSC, CMUG survey
	Reanalysis	10km	1 day	5%	5%	1%/decade	SSEOB	BSC, CMUG survey
	trend monitoring	20km	1 month	10cm or 10%	10cm or 10%	2 mm/decade	SSEOB	BSC, CMUG survey

	decadal f/c	50km	1 month	10cm or 10%	10cm or 10%	2 mm/decade	SSEOB	BSC, CMUG survey
Sea-ice thickness	Initialise	20km	1 day	10cm or 10%	10cm or 10%	2 mm/decade	SSEOB	BSC, CMUG survey
	Reanalysis	10km	1 day	10cm or 10%	10cm or 10%	2 mm/decade	SSEOB	BSC, CMUG survey
	trend monitoring	10km	1-2-7 day	0.01 m/s	0.01 m/s	0.01 m/s/decade	SSEOB	BSC, CMUG survey
Sea-ice drift	Initialise	5km	1 day	0.01 m/s	0.01 m/s	0.01 m/s/decade	SSEOB	BSC, CMUG survey
	Reanalysis	10km	1 day	0.01 m/s	0.01 m/s	0.01 m/s/decade	SSEOB	BSC, CMUG survey
	trend monitoring	10km	1-2-7 day	2%	5%	1%/decade	SSEOB	BSC, CMUG survey
Melt pond fraction	Initialise	5km	1 day	2%	5%	1%/decade	SSEOB	BSC, CMUG survey
	Reanalysis	10km	1 day	2%	5%	1%/decade	SSEOB	BSC, CMUG survey

Table 17. Requirements for satellite derived observations of sea-ice

4.16 Clouds

The IPCC AR5 report states that clouds and aerosols continue to contribute the largest uncertainty to estimates and interpretations of the Earth's changing energy budget. Progress has been made in the understanding of how cloudiness and humidity changes simulated by climate models in warmer climates are related to large-scale circulation changes, such as the rising of high clouds and poleward shift of clouds associated with the storm tracks. However, some of the cloud changes vary substantially among models and are likely due to sub-grid scale processes, including the representation of convection and aerosol-cloud interactions in models. The uncertainty in the sign and magnitude of the cloud feedback in CMIP3 and CMIP5 models is dominated by uncertainties in the response of low-level clouds in tropical and subtropical regions to warming (Boucher *et al*., 2013). In CMIP6 models, also changes in clouds over the Southern Ocean can play an important role (Zelinka *et al*., 2020). Likewise, errors in the representation of model clouds have been invoked to explain important and long-standing mean-state climate model biases, like the substantial Southern Ocean warm biases present in most CMIP5 climate models (Hyder et al 2018).

The use of satellite data has increased since IPCC AR5, due to data records reaching a useful length and there is more available data from passive and active sensors as well as new types of technologies. The WCRP Grand Challenge on Clouds, Circulation and Climate Sensitivity is focused around five main initiatives, the fourth one, "Leveraging the past record", aims to exploit observations of the recent past, or proxies for longer-term changes, to better constrain cloud processes and feedbacks[17]. The Cloud-CCI data-set currently covering 32 years contributes to addressing this challenge by adding a new data set with consistent cloud variables and uncertainty information.

[17] http://www.wcrp-climate.org/grand-challenges/gc-clouds

The GEWEX Cloud Assessment coordinated intercomparison of L3 cloud products of 12 global "state of the art" datasets has shown how cloud properties are perceived by different instruments and how cloud property averages and distributions are affected by instrument choice as well as some methodological decisions (Stubenrauch *et al*., 2013). In the assessment they found that differences in long-term variation in global-mean cloud amount between the datasets were comparable in magnitude to the interannual variability (2.5-3%). Still, these satellite cloud products are very valuable for climate studies or model evaluation, the geographical and seasonal variations in the cloud properties agree very well. They do not agree as well over deserts and snow-covered regions and for high level cloud statistics, due to problems detecting thin cirrus (Stubenrauch *et al*., 2013).

For process studies there is also a strong requirement for satellite observations to improve the representation of clouds in climate models and here the long term stability is not an important requirement, as the data are used to investigate changes on timescales of hours to seasons.

When comparing to climate models, observation time and view from above as well as retrieval filtering have to be taken into account. This can be achieved either by simple methods or by using observation simulators for the different datasets as in the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package (Bodas-Salcedo *et al*., 2011), which consists of individual simulators, with each corresponding to a specific cloud dataset (e.g., ISCCP, CALIPSO, MODIS, MISR, or CloudSat). Cloud-CCI efforts to develop a Cloud-CCI simulator, as well as testing more simple methods to be used by models without all fields available, follow the GEWEX Clouds Assessments recommendations.

The main uses of the Cloud-CCI datasets by the CMUG survey participants range from comparisons with models, for improved process understanding and parameterisations to detecting climate trends on regional and seasonal scales. The major obstacles expressed in using satellite data are concerns about drifts and continuity between satellites and platform and lack of documentation.

Here, we first discuss the precise requirements on horizontal and vertical resolution, observing cycle and the type of usage as summarized in Table 18. Thereafter the more general requirements and comments from the survey are given.

Horizontal resolution

Current global climate models are run typically sub-100 km and regional climate and NWP models have resolutions from 50 km down to km scale. For detailed process studies it is desirable to have information at sub-grid scales, hence the specification of 10 km and for high resolution models km scale. For more general evaluation studies, e.g. comparison of monthly mean geographical distributions, this could be relaxed considerably and horizontal resolutions of around 50 km can still be useful.

Vertical resolution

The distribution of the vertical levels in atmospheric models is highly non-linear with respect to altitude – the layers are typically much more tightly spaced in the boundary layer compared to the free troposphere, for example. Current global climate models have vertical resolutions of around 200 m in the boundary layer (with even this not being entirely satisfactory to represent stratocumulus clouds), increasing to around 500 m in the middle troposphere – the specification of 100 m is thus again based on the requirement for process studies. This could also be relaxed for other evaluation work and a vertical resolution of 500 m (or more) might be useful, depending on the information content of the particular observations. Vertically resolved clouds from CloudSat and CALIPSO have been used extensively over the last ten years by the CMC. For the passive sensors used for the Cloud-CCI products there is no vertical information, except cloud top height and cloud top pressure. For validation purposes it would be useful to have these products compared to the CloudSat/CALIPSO data sets.

Observing cycle

In common with many related processes (e.g. rainfall, convection) the diurnal cycle of cloud remains a common weakness in the majority of current models. Examples of cloud systems with large diurnal cycles are tropical convection over land and marine stratocumulus clouds. Ideally, data with a temporal resolution comparable to the typical model time step (15-30 minutes) would be desirable. Again, however, much useful information could be obtained with 1-hourly data, with the upper limit on utility probably being 2-3 hours.

Model development/evaluation

There are various products of interest which range from fields of cloud cover and top pressure/temperature to profiles of cloud water and ice concentration. The utility of statistical summaries (e.g. optical depth vs cloud top pressure histograms) when employing the COSP simulator, can be compared to climate model output in a very straightforward manner. This has been recognised by the observational community and ISCCP-like histograms are now produced using both MODIS and MISR data. This approach has several advantages:

- It puts the CCI data into a format that is already familiar to modellers.
- It allows the CCI data to be easily compared to other cloud data sets.
- It allows the CCI data to be easily integrated into pre-existing and tested methods for exploiting satellite cloud data for model evaluation.

Parameter	*Application*	*Horizontal Resolution*	*Vertical resolution*	*Observing Cycle*	*Precision*	*Accuracy*	*Stability*	**Error Type (see Table 3)**	**Source**
Cloud cover	*model development*	*50km-1km*	*N/A*	*Monthly to 1h*	*10%*	*5%*	*1%/year*	*SSEOB*	CMUG survey
	trend monitoring	*50km*	*N/A*	*Monthly to 3h*	*10%*	*5%*	*1%/decade*	*SSEOB*	CMUG survey
	Reanalysis/ Processes	*10km-2km*	*N/A*	*6h to1h*	*10%*	*10%*	*1%/year*	*SSEOB*	CMUG survey
Cloud top height	*model development*	*10km*	*N/A*	*Monthly to1h*	*0.1km*	*0.1km*	*0.1km/ year*	*SSEOB*	CMUG survey
	data assimilation	*5km*	*N/A*	*1h*	*0.1km*	*0.1km*	*N/A*	*ERRCOV*	CMUG survey
	trend monitoring	*30km*	*N/A*	*Monthly to 3h*	*0.2km*	*0.2km*	*0.1km/ decade*	*SSEOB*	CMUG survey
	Reanalysis/ Processes	*10km-2km*	*N/A*	*6h to 1h*	*10%*	*0.1km*	*0.1km/year*	*SSEOB/ ERRCOV*	CMUG survey
Cloud top temp	*model development*	*10km*	*N/A*	*1h*	*0.1K*	*0.25K*	*<0.1km/year*	*SSEOB*	CMUG survey
	trend monitoring	*30km*	*N/A*	*3h*	*0.25K*	*0.25K*	*0.25K/decade*	*SSEOB*	CMUG survey
Cloud ice profile	*model development*	*50km-1km*	*0.2km*	*1h*				*SSEOB*	CMUG survey

Cloud water profile (> 100 µm)	*model development*	*50km-1km*	*0.2km*	*1h*				*SSEOB*	CMUG survey
Cloud water profile (< 100 µm)	*model development*	*50km-1km*	*0.2km*	*1h*				*SSEOB*	CMUG survey
Cloud effective radius?	*model development*	*50km-1km*	*0.2km*	*1h*	*1um*	*1um*	*1um*	*SSEOB*	CMUG survey

Table 18. Requirements for satellite derived cloud observation. The underlined values for the horizontal resolution are requirements from regional climate modellers.

Trend monitoring
The requirements for trend detection are somewhat more difficult to ascertain. Firstly, there is currently no clear indication from presently available observations about cloud trends and secondly this may well be too stringent a test for current models, given the known uncertainties in the representation of cloud processes. It certainly is the case that the cloud modelling and cloud feedback community is currently much more focused on process studies than on long-term trends. That said, a new data set that was able to determine trends in cloud amount, for example, with the specified level of accuracy/stability would be a major advance and would undoubtedly be of great interest to climate modellers.

The GCOS requirements for the cloud ECV are somewhat relaxed in terms of observing cycle (3-6hr) compared to the CMC requirements which may reflect the needs in terms of long-term trend monitoring rather than model process studies. Also the GCOS accuracies for cloud cover and cloud top height are more relaxed than those required for model processes.

Single sensor vs merged sensor products
Another consideration is that the generation of merged products from quite different sensors will be difficult to interpret for most applications. Such merged products are difficult to use indeed, the rationale behind the simulator approach is precisely to avoid such difficulties by generating model equivalents of single-sensor products. However, the CMC are interested in both single sensor and merged products. To ensure traceability for merged products it is important to provide pixel (grid-point) uncertainties and good documentation of the processing for merged products.

Satellite data validation/evaluation, format and access
For validation, the CMC users recommend it should be done for all seasons, for day and night and on regional scale against reference data with known errors, e.g. against station data and in-situ measurements and CloudSat/CALIPSO. Scores could be combined addressing bias, spatial and temporal correlations. The preferred format of the data is netCDF. Many modellers say it would be very useful to follow the CMIP5/6 format (CMOR standard), and some even say that it is a prerequisite for extensive use of the data within the climate community. The preferred means of access to the data is via ftp or via a web browser. Some strongly recommend that the data is available from a centralised server as ESGF and/or that the data is available through obs4MIPs.

Finally, to summarize, the general view on cloud satellite data from AR5, recent papers and the participants in the user survey lead CMUG to recommend that the cloud ECV datasets are continued to be designed for validating cloud model processes as well as building a long term monitoring datasets, despite difficulties. A simulator should accompany the data and the data should have been validated and include uncertainties and be well documented. For physical cloud parameters such as liquid water path or ice water path, it is important to distinguish between in-cloud values and grid-box averages. CMIP models typically provide grid-box averages only.

4.17 Ozone

The ozone concentration in the atmosphere (mainly the total ozone column) has been measured for several decades since the late 1980s when the impact of human activities on the upper stratosphere and lower stratosphere chemical processes was discovered, resulting in the high latitude ozone holes. Monitoring the trends of ozone content remains a key issue for the study of the recovery of stratospheric ozone and also for monitoring human induced greenhouse gases as far as tropospheric ozone is concerned. It is also essential to study stratospheric-tropospheric exchange processes and to give a better representation of the dynamics, chemical, transport and radiative processes. Ozone data assimilation is of primary importance for environmental studies including the initialization of air quality prediction (interactions between air quality and climate are deemed increasingly important). Some studies have also revealed the potential of ozone observations in constraining the atmospheric dynamics through data assimilation. Considering available observations, those from satellites are crucial in providing information on the ozone content of the atmospheric column but also, through the development of new sensors, to provide valuable information on partial columns and also the ozone profile.

The requirements for the ozone ECV are given in Table 19.

Parameter	***Application***	***Horizontal Resolution (km)***	***Vertical Resolution (km)***	***Observing Cycle (h)***	***Precision (%)***	***Accuracy (%)***	***Stability (%)***	**Error Type (see Table 3)**	**Source**
Ozone profile									
Higher stratosphere & mesosphere (HS & M)	*Model Development and Evaluation*	*100*	*3*	*24*	*5*	*10%*	*2.0 %/decade*	*SSEOB*	CMUG survey
	Reanalysis and Data Assimilation	*50*	*1*	*6*	*5*	*10%*	*2.0 %/decade*	*SSEOB*	CMUG survey
Lower stratosphere (LS)	*Model Development and Evaluation*	*50*	*2*	*24*	*3*	*6%*	*2.0 %/decade*	*SSEOB*	CMUG survey
	Reanalysis and Data Assimilation	*20*	*1*	*6*	*3*	*6%*	*2.0 %/decade*	*SSEOB*	CMUG survey
Higher troposphere (HT)	*Model Development and Evaluation*	*20*	*2*	*24*	*3*	*8%*	*2.0 %/decade*	*SSEOB*	CMUG survey
	Reanalysis and Data Assimilation	*20*	*1*	*6*	*3*	*6%*	*2.0 %/decade*	*SSEOB*	CMUG survey

Lower troposphere (LT)	*Model Development & Evaluation*	*20*	*2*	*24*	*6*	*10%*	*2.0 %/decade*	*SSEOB*	CMUG survey
	Reanalysis and Data Assimilation	*20*	*1*	*4*	*5*	*10%*	*2.0 %/decade*	*SSEOB*	CMUG survey
Ozone column									
Troposphere column	*Model Development and Evaluation*	*20*		*24*	*6*	*15*	*2.0 %/decade*	*SSEOB*	CMUG survey
	Reanalysis and Data Assimilation	*20*		*4*	*5*	*10*	*2.0 %/decade*	*SSEOB*	CMUG survey
Total column	*Model Development and Evaluation*	*20*		*24*	*2*	*4*	*1.0 %/decade*	*SSEOB*	CMUG survey
	Reanalysis and Data Assimilation	*20*		*6*	*3*	*5*	*1.0 %/decade*	*SSEOB*	CMUG survey

Table 19. Requirements for satellite derived observation of ozone.

For lower troposphere and the tropospheric column, CMUG prefer more stringent requirements than those stated in Table 20 on the observing cycle to better constrain O_3 pollution episodes and the daily cycle. For the latter, hourly observations would be needed. This type of frequency can only be achieved with sensors on board of geostationary satellites, such as the proposed Sentinel-4 and TEMPO missions.

As far as ozone assimilation is concerned, products from single sensors would be preferred to merged products. Merged products if they are all obtained with the same technique and over a long period span (like the SBUV sensors over 30 years) are useful in a model validation context like CCMVal (Eyring *et al*., 2010), aiming at evaluating each process separately. This implies to provide these different products as separate datasets.

The error/uncertainty requirements have become generally more stringent for precision while the requirements in terms of accuracy have been slightly relaxed. This could be a consequence of the fact that many models nowadays include schemes to correct the observations for systematic biases, of which the accuracy is an estimate.

User friendly quality information and traceability have been identified as one of the major obstacles in current satellite data usage. While good documentation, especially on the quality assessment, and history of changes (with appropriate data versioning) are also regarded as important aspects to efficiently use the data.

A homogenous and coherent definition of the tropopause (possibly also included in the dataset) was suggested as being very important and useful for some applications.

4.18 Greenhouse Gases

A comprehensive understanding of greenhouse gases is crucial for informing societal response to climate change. Applications with a need for observations of greenhouse gases such as CO_2 and CH_4 include Model Development, Decadal Forecasting and Regional Source/Sink Determination. As shown in Table 20, each application has somewhat different observational requirements reflecting the particular aspect of greenhouse gases under consideration.

The principal products for GHG observational requirements for Regional Source/Sink Determination are:

- 4-dimensional gridded fields of CO_2 and CH_4 produced in near-real-time (based on data assimilation of near-real-time data products, typically from operational satellites),
- 4-dimensional gridded fields of CO_2 and CH_4 produced in "delayed mode" (6 months delay, to allow data assimilation of research-mode satellite data products and reanalysis applications),
- 3-dimensional gridded fluxes of CO_2 and CH_4 produced in "delayed mode",
- Re-analysed concentration and flux fields of CO_2 and CH_4 for the last decade.

Flux fields are an important factor for decision-makers at several levels and need to be estimated with confidence. The fidelity of flux estimates is strongly influenced by accuracy and stability of the observations that are used as input to the data assimilation and re-analysis systems. This drives the requirements given in Table 20 for some of the required parameters.

The data outputs of this CCI+ ECV project are of interest to the budget closure project on the carbon cycle (RECCAP2).

Parameter	Application	Horizontal Resolution	Vertical Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Trace gas profile CH_4 - Troposphere column	Regional source/sink determination	5/20/50 km	N/A	3/4/6 h	0.1/0.5/1% 2/10/20 ppb	0.1/0.5/2.0 % 2/10/20 ppb	0.5/0.7/2.0 %/dec 2/7/35 ppb/dec	SSEOB	ECMWF/ CMUG survey
Trace gas profile CH_4 - Total column	model development	25km	N/A	6 h	1%	1%	10ppb/dec	SSEOB	ECMWF/ CMUG survey
	decadal f/c	20km	N/A	Daily	<<10 ppb	<<10 ppb	2%/dec 35 ppb/dec	SSEOB	ECMWF/ SMUG survey
	Regional source/sink determination	10/50/100 km	N/A	3/4/6 h	0.25/0.5/1% 5/10/20 ppb	0.1/0.5/2.0 % 2/10/40 ppb	0.1/0.5/2.0 %/dec 2/10/35 ppb/dec	SSEOB	ECMWF/ CMUG survey
Trace gas profile CO_2 - Total column	model development	25km	N/A	6h	0.5/1ppm	0.5/1ppm	0.1/0.5ppm /dec	SSEOB	ECMWF/ CMUG survey
	decadal f/c	2/5/20km	N/A	Daily	0.3/0.5/1% 1/1.5/3 ppm	0.3/0.5/1% 1/1.5/3.0 ppm	0.5/1.5/2 %/dec 2/5/8 ppm/dec	SSEOB	ECMWF/ CMUG survey
	Regional source/sink determination	5/20/50 km	N/A	3/6/24 h	0.25/0.5/0.7 5% 1/2/3 ppm	0.25/0.5/1% 1/2/4.0 ppm	0.5/1.5/2 %/dec 2/5/8 ppm/dec	SSEOB	ECMWF/ CMUG survey

Trace gas profile CO_2 - Troposphere column	Regional source/sink determination	5/20/50 km	N/A	3/4/6 h	0.15/0.4/0.5 % 0.5/1.5/2 ppm	0.15/0.5/1% 0.5/1.5/4.0 ppm	0.15/0.5/2 %/dec 0.5/1.5/7.5 ppm/dec	SSEOB	ECMWF/ CMUG survey

Table 20. Requirements for satellite derived observation of greenhouse gases.

The requirements are given for tropospheric and total column only, in recognition that requirements for profile data would be very demanding for existing satellite data. In the event that data providers consider it feasible to provide profile data approaching GCOS requirements, then more refined user requirements could be given in a future update of this document. It is also important that user requirements should not be based on what is currently possible, but rather on what is needed. So, if GHG profiles are needed, requirements should be set for those. A gap analysis could then establish whether current satellite instruments can meet or not the given requirements, The user community increasingly asks for horizontal and vertical resolution in the Lower Stratosphere to be the same as that for the Higher Troposphere, in contrast to previous GCOS requirements. Other applications of greenhouse gas observations may have different sets of requirements, for example, the detection of CH_4 emissions from pipelines or similar small sources would require higher horizontal resolution and vertical resolution in the lower troposphere.

Turning to the GHG observation requirements for decadal and climate prediction it is principally the distribution of the trace gases at the start of the forecast that can be important to help define the atmospheric fields, and for this a high observing cycle is useful. Additionally, more stringent requirements than described in Table 21 have been made for the horizontal resolution that is now comparable with that needed in other applications.

Similar to the ozone section above, it would be important to provide not only merged GHG products but also products from single sensors as separate datasets. Users also pointed out that the harmonisation between the various datasets is a key aspect to efficiently using the data.

4.19 Aerosols

The impact of aerosols on climate is an important factor governing climate change. Aerosols have offset part of the warming expected from anthropogenic emissions of greenhouse gases but with decreasing emissions of sulphate aerosols which have a cooling effect and increaseing emissions of black carbon (also from forest fires) which has a warming effect, this might be different in the future. It is very important to decrease the uncertainties on the aerosol forcing because this will contribute to better constrain the climate sensitivity from current observational climate records. As a result measurements of atmospheric aerosols (both tropospheric and stratospheric) are required. There is a further arbitrary split at 3 km height to obtain aerosol products below and above the lower troposphere. This split is somewhat arbitrary and aerosol profile information from spaceborne lidars should be instead considered.

Aside from the direct radiative effect it is in particular the impact of indirect radiative effects (mainly through aerosol-cloud interactions) which needs to be better understood to better estimate the climate sensitivity to aerosols in climate models. Thus, there are two aspects that need to be addressed. Relatively high resolution data with associated environmental data (e.g. clouds) for a better process understanding, as well as long-term monitoring on global scales to address trends in aerosol properties. Precipitation has also been reported by the aerosol product users as an important parameter to constrain aerosol loads since it is the main mechanism of removal.

The parameters for aerosol climatologies within global models are listed in Table 21. It includes the aerosol extinction optical depth (AOD) (at the modelling reference wavelength at 550nm) for both the total atmospheric column as well as stratified over four atmospheric altitude sections to distinguish between stratosphere (important after major volcanic eruptions) and tropospheric layers linked to high-, mid- and low level clouds. It is noted that this requirements related to these four altitude sections might become obsolete when profiling data from lidars are considered. Upper tropospheric aerosols have enhanced capabilities for long range transport, while lower tropospheric aerosols remain more local and influence the near surface meteorology (e.g. visibility, air quality). In general, tropospheric AOD can be derived as the total AOD minus the stratospheric AOD. In addition to total extinction optical depth (absorption + scattering) the absorption optical depth is also an important parameter to measure and has more stringent accuracy requirements being only part of the total extinction.

Aside from aerosol amount also the aerosol composition is of interest. A very useful property in that sense is data for AOD at different wavelengths. These different AOD data provide information on aerosol size. AODs at two different wavelengths already define the Angstrom parameter, which is a more general size-indicator. Even better is the AOD fine mode fraction, which requires AOD data at least four different wavelengths in the visible and the near-IR. Then via the Angstrom parameter spectral dependence the total AODs can be stratified into fractions associated with smaller (radii <0.5µm) and larger sizes (radii >0.5µm). Thus, aside from the AOD retrieval at 0.55µm, additional AOD retrievals at one or even better at three other wavelengths in the visible or near-IR are desirable (e.g. 443nm, 670nm, 870nm). The 412 nm wavelength could also be useful to characterise dust aerosols (I.e. Deep Blue product from MODIS), Other useful elements to characterize aerosol type are data on polarization and absorption. Polarization provides information on aerosol shape (e.g. mainly to discriminate dust from other aerosol types). In most retrievals *a-priori* assumptions on aerosol absorption properties are made and these must be provided in the metadata associated with the data set.

Concerning single sensor datasets or merged product datasets or both, the CMC stressed to that both are required. Merged (single sensor) products are preferred for monthly mean (instantaneous) data. Traceability back to the sensor and documentation are important issues. Moreover, merged products are more difficult to achieve in NRT (Near Real Time) and that’s why they are not a preferred option for aerosol analyses. However, in climate and reanalysis applications they can have an important role, particularly if the systematic and random errors on the merged product are well characterized.

One additional CMC requirement is defined by the assessment of aerosol processes in climate models which requires data on associated environmental properties and the potential interactions with clouds. Thus, data on clouds (from the cloud ECV) are required which match in terms of spatial and temporal) resolution, observing period and if possible satellite platform.

Concerning the preferred validation methodology, the CMC stressed cross-validation against in-situ data, ground-based measurement (e.g. AERONET), other instruments (e.g., lidar data from CALIPSO, EARLINET and other lidar networks) and reanalysis data.

The requirements given in Table 21 are likely to change in the next few years as the GCOS ECV parameters for aerosols are being reviewed. There is consensus in the community that this optical depth below and above 3km is an arbitrary split providing no benefit. Requirements regarding the aerosol vertical structure will become more prominent, particularly following the successful CALIPSO mission and lidar technology from space becoming more reliable for aerosol applications.

Parameter	Application	Horizontal Resolution	Observing cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Total extinction optical depth (at 4 VIS + IR wavelengths)	model development	1km	1hr	0.02	0.02	0.02/decade	SSEOB	ECMWF/CMUG survey
	assimilation	2km	1hr	0.02	0.02	0.02/decade	SSEOB	ECMWF/CMUG survey
	decadal f/c	2km	Daily	0.01	0.02	0.005/decade	SSEOB	ECMWF/CMUG survey
	trend monitoring	2km	3hr	0.005/ 0.01	0.01/ 0.02	0.02/decade	SSEOB	ECMWF/CMUG survey
Total aerosol absorption optical depth at 0.55um	model development	1km	1hr	0.004	<0.01	0.005/decade	SSEOB	ECMWF/CMUG survey
	trend monitoring	2km	3hr	0.002/ 0.01	0.004/ 0.02	0.002/ decade	SSEOB	ECMWF/CMUG survey
Aerosol optical depth in stratosphere (at 4 VIS + IR wavelengths)	model development	1km	1hr	0.02	0.02	0.02/decade	SSEOB	ECMWF/CMUG survey
	trend monitoring	2km	6hr	0.02	0.02	0.01/decade	SSEOB	ECMWF/CMUG survey
Aerosol optical depth in troposphere (at 4 VIS + IR wavelengths)	model development	1km	1hr	0.004	0.02	0.02/decade	SSEOB	ECMWF/CMUG survey
	trend monitoring	2km	6hr	0.002	0.004	0.01/decade	SSEOB	ECMWF/CMUG survey
Aerosol optical depth above ~3km (680hPa) (at 4 VIS + IR wavelengths)	model development	1km	1hr	0.01	0.02	0.02/decade	SSEOB	ECMWF/CMUG survey
	trend monitoring	2km	6hr	0.005	0.01	0.01/decade	SSEOB	ECMWF/CMUG survey
Aerosol optical depth below ~3km (680hPa) (at 4 VIS + IR wavelengths)	model development	1km	1hr	0.01	0.02	0.02/decade	SSEOB	ECMWF/CMUG survey
	trend monitoring	2km	6hr	0.005	0.001	0.01/decade	SSEOB	ECMWF/CMUG survey
Aerosol lidar depolarisaton ratio (VIS)	model development	1km	1hr	N/A	10%	N/A	SSEOB	ECMWF/CMUG survey
	trend monitoring	2km	6hr	N/A	5%	N/A	SSEOB	ECMWF/CMUG survey

Table 21. Requirements for satellite derived aerosol datasets.

4.20 Glaciers

Glaciers provide a visible indication of the effects of climate change, as the mass balance at the surface of a glacier (the gain or loss of snow and ice over a hydrological cycle) is determined by the climate. It is important to measure and understand the areal and volumetric changes with time, and also how well climate models can represent or parameterise glaciers.

According to the tiered strategy of global glacier monitoring in the Global Terrestrial Network for Glaciers (GTN-G), the basic application of satellite data is the generation of repeat glacier inventories at decadal time scales using semi-automated classification techniques and data processing in Geographic Information Systems. This is in line with GCOS requests for a globally complete map of glaciers. The global map of glaciers would serve several fields of application, including:

- improved modelling of global sea-level rise,
- a sound basis for change assessment,
- an important input for hydrological and glaciological modelling.
- a possibility to validate output from Regional Climate Models (RCMs), and
- a data set to initialise the land ice fields in RCMs.

Apart from application in defining glacier extent, satellite data are used widely to derive further glaciological parameters including snow facies, velocity fields and elevation changes. All these products do strongly vary in terms of sensors (resolution), observing period and cycle, or required precision and accuracy. A list of satellite based observational requirements and capabilities was compiled by IGOS (2007), IGOS was dissolved in 2008 and a WMO Global Cryosphere Watch[18] (GCW) project set up to continue its work. Both IGOS (2007) and the GCW website were used as a base for Table 23 below. The long term stability of the measurements is crucial for this ECV as it is an indicator of climate change.

The two main requirements for the glacier datasets for the CMC are trend monitoring and providing initial conditions for climate models. For both of these the requirements described for GTN-G are more than sufficient. The datasets can also be used for validation of land surface process in climate model predictions which have the same requirements for accuracy as the trend monitoring.
This ECV is of interest to the IMBIE project.

Parameter	Application	Horizontal Resolution	Observing Cycle	Precision	Accuracy	Stability	Error Type (see Table 3)	Source
Glacier Area	Initialisation	15 m	1 year	0.01km^2	<1%		SSEOB	IGOS 2007/GCW
	trend monitoring	30 m	5 years	0.01km^2	<1%	0.01km^2/ decade	SSEOB	IGOS 2007/GCW
Glacier Topography	Initialisation	<100 m	1 year	1 m	5 m		SSEOB	IGOS 2007/GCW

[18] https://globalcryospherewatch.org/reference/obs_requirements.php

	trend monitoring	<100 m	5-10 years	1 m	5 m	1 m/ decade	SSEOB	IGOS 2007/GCW
Velocity	Initialisation	20 m	1-12 months	1 m/yr	10 m/yr		SSEOB	IGOS 2007/GCW
	trend monitoring	20 m	1 year	1 m/yr	10 m/yr	1 m/ decade	SSEOB	IGOS 2007/GCW
Snowline	Initialisation	30 m	1 year	30 m	100 m		SSEOB	IGOS 2007/GCW
	trend monitoring	30 m	1 week / 1 year	30 m	100 m	30 m / decade	SSEOB	IGOS 2007/GCW

Table 22. Requirements for satellite derived Glacier observations.

4.21 Fire

Fire disturbances alter vegetation dynamics and impact climate. Climate models that account dynamically for climate induced changes in vegetation simulate fire disturbance within process-based fire sub-models. The development and evaluation of such sub-models depend on the availability and quality of satellite-based fire disturbance observations. Such complex Earth System models are crucial to assess fire climate interactions and the impact of fire on the global carbon cycle.

In addition, global vegetation models can be utilized to diagnostically simulate fire emissions by combining information on burned area, available fuel load and burning conditions. Satellite based burned area products can thereby serve both as prescribed boundary conditions or as alternative verification references. Besides uncertainties in burned area estimates, such an approach is limited by an uncertain quantification of available fuel loads and burning conditions (e.g. combustion completeness, mortality rates, emission factors). Fire disturbance products will therefore be best exploited in models when consistently derived ancillary data products, such as land cover classification or biomass availability, are provided that help to constrain specific burning conditions.

The assessment of fire emissions will be one important application of fire disturbance products. Fire emissions serve as boundary conditions for atmospheric aerosol and chemistry models used to assess air quality and/or the influence of atmospheric chemistry on the climate. An operational usage of atmospheric composition models will require near real-time availability of the fire disturbance ECV. Other applications of the fire CCI product include improvement of fire model parameterisations and process studies.

The strong interannual variability of fire activity will require data products that cover a multiyear timespan for the development and evaluation of process-based fire models as well as for the application of satellite observed burned area products as boundary conditions. The current CCI Fire product extends for 18 years, and thus meets this requirement.

The specific requirements for the fire disturbance ECV are listed in Table 23. In terms of spatial resolution and observing cycle these are close to the GCOS requirements.

Parameter	Application	Horizontal Resolution	Observing Cycle	Accuracy	Stability	Source
Burnt area	**GCOS**	30m	24 hours	15% compared to 30 m observations		
	Trend monitoring/ verification	0.25/1.0/5.0 km	1/1.5/3 d	30/20/10 %(MAX)	5.00%	MPI-M
	Prescribe model boundary condition	0.25/1.0/5.0 km	3h/ 1/1.5/3 d/monthly	30/20/10/1 %(MAX)	5.00%	MPI-M

Table 23. Requirements for satellite derived observations of *burnt area parameters*

Users do apply level 1 (direct data assimilation), level 2 (assimilation), as well as level 3 data (verification and climate monitoring). The users want single datasets as well as merged datasets. For the merged product traceability to the single sensors as well as a good documentation is a key requirement. The validation should be based on site level data and inter-instrument comparison. The preferred data format is netCDF–CF following the CMIP6 format guidelines, made accessible via FTP. For climate applications, however, the product requirements for regional applications should be comparable with those for global applications. The data outputs of this CCI+ ECV project are also of interest to the budget closure project on the carbon cycle (RECCAP2).

No fire radiative power product is available from the fire CCI project. This is not strictly an ECV although it is a requirement of climate modellers.

4.22 Soil Moisture

Soil moisture is an important variable for all models from NWP to climate time scales. The GCOS (GCOS 2016) requirements are given in Table 24 below along with those required for modelling and NWP data assimilation systems. The GCOS requirements now include new parameters that were developed since the last CMUG User Requirements survey (2015)[19]. Table 25 then outlines the requirements for a wider range of climate monitoring and services.

Parameter	Application	Horizontal Resolution	Observing Cycle	Accuracy	Stability	Error Type (see Table 3)	Source
Volumetric soil moisture (up to 5cm depth)	Initialisation	1-10km	Daily	$0.035m^3/m^3$		SSEOB	MF/CMUG survey
	trend monitoring	1-20km	Daily	$0.04m^3/m^3$	$0.01m^3/m^3/yr$	SSEOB	MF/CMUG survey
	GCOS	1-25 km	Daily	$0.04\ m^3/m^3$	$0.01\ m^3/\ m^3/year$		MF/CMUG survey

[19] CMUG Phase 2 Deliverable 1.1: User Requirement Document (v0.6), available at http://ensembles-eu.metoffice.com/cmug/CMUG_PHASE_2_D1.1_Requirements_v0.6.pdf.

Freeze / thaw	GCOS	1-25 km	Daily	90%		Flag indicating whether the land surface is frozen or not	MF/CMUG survey
Surface inundation	GCOS	1-25 km	Daily	90%		Flag indicating whether the land surface is inundated or not	MF/CMUG survey
Root zone soil moisture	GCOS	1-25 km	Daily	0.04 m^3/m^3	0.01 m^3/m^3/year	Average soil moisture content in root-zone layer	MF/CMUG survey

Table 24. GCOS and modelling requirements for soil moisture observations.

Soil moisture is widely used to initialise surface fields in models and is of particular importance for seasonal climate predictions and the monitoring of moisture anomalies on the terrestrial land surface. There is strong need for consistency in this ECV with other ECVs for example temperature, surface humidity, albedo, vegetation (AGB, LAI, FAPAR) and precipitation. No differences in requirements between global and regional models were expressed in the survey.

According to the CMUG user survey, soil moisture observations are essential in all application domains for climate modelling. The widest expected use of soil moisture data is in the field of model development (process studies) and model evaluation. 100% of the expert users are interested in using soil moisture data for these applications, while 60% use it also for model initialization and climate monitoring and attribution.

The detailed requirements for ECV soil moisture collected by CMUG are summarized in Table 25. The requirements for the observation cycle are that most users and applications require daily data. Monthly data might be sufficient for some applications like e.g. trend monitoring, while even higher (sub-daily) temporal resolution would be desired for special process studies.

The CMC support target and threshold values as defined by the EUMETSAT HSAF as correlation coefficient values with ground-truth observations of R = 0.65 and R = 0.50, respectively. Accuracy target is RMSE of 0.04 m^3m^{-3} for SM $\leq$ 0.2 m^3m^{-3} and better than 20%, i.e., RMSD/mean in %, for SM higher than 0.2 m^3m^{-3}. More stringent requirements might apply on larger spatial scales due to the spatial aggregation of potential errors. In addition several users emphasized the need for information on the depth where the soil moisture data is sensed.

One criterion for soil moisture records was for temporal stability in long-term data. Overall, the most important aspect for the users is that the datasets show a long term stability without sudden jumps or data gaps. A quantitative accuracy measure was not given and is therefore not provided in the summary table.

For error measures, in case of individual sensor measurements, the uncertainty on the single sensor retrievals should be provided, while for L3 data the uncertainty of the merged product is needed. The latter requires an uncertainty model to quantify adequately uncertainties from spatial upscaling/regridding procedures as well as effects of spatiotemporal sampling patterns on random and systematic error components.

Parameter	**Application**	**Horizontal Resolution**	**Observing Cycle**	**Precision**	**Accuracy**	**Stability**	Error Type (see Table 3)	Source
Soil Moisture								
Volumetric SM	trend monitoring	< 1 km² to 25x25 km²	Daily … monthly	0.005 – 0.01 [m³/m³]	0.5vol.%(SH)	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France
	model initialisation / boundary condition	< 1 km² to 25x25 km²	Daily	0.005 - 0.035 [m³/m³]	1% / 0.5% (SH) Larger deviations are of less concerns than for temporal anomalies (also strong spatial variability)	No information available	SSECDR ERRMERG ERRCOV	CMUG survey/ Meteo-France
	Validation	< 1 km² to 25x25 km²	Daily	0.005 - 0.035 [m³/m³]	-	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France
	Monitoring/ Attribution	< 1 km² to 25x25 km²	Daily	0.005 - 0.035 [m³/m³]	-	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France
	Data assimilation	< 1 km² to 25x25 km²	Daily	0.005 - 0.035 [m³/m³]	0.04 [m³/m³]	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France
	Model development and validation	< 1 km² to 50x50 km²	Daily	0.005 - 0.035 [m³/m³]	0.04 [m³/m³]	No information available	ERRMERG	CMUG survey/ Meteo-France
Volumetric SM temporal anomalies (removing long term mean)	trend monitoring	< 1 km² to 25x25 km²	Daily	No information available	Larger deviations are of less concerns than for temporal anomalies (also strong spatial variability)	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France
	Prescribe model boundary condition	< 1 km² to 25x25 km²	Daily	0.005 - 0.035 [m³/m³]	Known & constant/	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France
Soil moisture anomalies	trend monitoring	< 1 km² to 25x25 km²	Daily … monthly	min{0.04 [m³/m³]; 10% relative of anomaly}	min{0.04 [m³/m³]; 5% relative of anomaly}	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France
	Prescribe model boundary condition	< 1 km² to 25x25 km²	Daily	min{0.04 [m³/m³]; 10% relative of anomaly}	min{0.04 [m³/m³]; 10% relative of anomaly}	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France
Profile soil moisture proxy	trend monitoring	< 1 km² to 25x25 km²	Daily … monthly	1 mm over rooting depth	1 mm	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France
	Prescribe model boundary condition	< 1 km² to 25x25 km²	Daily … monthly	1 mm over rooting depth	1 mm	No information available	SSECDR ERRMERG	CMUG survey/ Meteo-France

Table 25: Summary of user requirements for soil moisture ECV.

4.23 Total Anthropogenic Water Use

Defined as the volume of water used by country, by sector (agricultural, industrial and domestic)

Parameter	**Application**	**Temporal Resolution G/B/T**	**Observing cycle**	**Length of record**	**Error in measured value**	**Stability**	**Error Type (see Table 3)**	**Source**
Total Anthropogenic water use (Gm^3y^{-1})	GCOS 2022	1/-/12 Months	year		10/-/20 %			GCOS

Table 26: Summary of user requirements for Total Anthropogenic Water Use ECV

4.24 Vegetation parameters: FAPAR and LAI

The GCOS definition of FAPAR is the fraction of photosynthetically active radiation (PAR, i.e. the solar radiation reaching the surface in the 0.4-0.7µm spectral region) that is absorbed by vegetation canopy. Both black-sky (assuming only direct radiation) and white-sky (assuming that all the incoming radiation is in the form of isotropic diffuse radiation) FAPAR values may be considered. Similarly, FAPAR can also be angularly integrated or instantaneous (i.e., at the actual sun position of measurement). Leaves-only FAPAR refers to the fraction of PAR radiation absorbed by live leaves only, i.e., contributing to the photosynthetic activity within leaf cells.

Leaf Area Index of a plant canopy or ecosystem is defined as one half of the total green leaf area per unit horizontal ground surface area and measures the area of leaf material present in the specified environment (projection to the underlying ground along the normal to the slope). Effective Leaf Area Index is the LAI value that would produce the same indirect ground measurement as that observed assuming foliage distribution (LAIeff=LAItrue × canopy clumping index). The conversion of data measurements to true values is an essential step and requires additional information about the structure and architecture of the canopy, e.g. gap size distributions, at the appropriate spatial resolutions. Leaf Area Index controls important mass and energy exchange processes, such as radiation and rain interception, as well as photosynthesis and respiration, which couple vegetation to the climate system.

For modelling purposes requirements are

- a true LAI dataset, or at the very least derivation of the clumping index in association with the effective LAI dataset
- a common de-noised product
- guidelines for the use of data quality flags supplied with the data

CMUG have been closely involved in the preliminary stages of the vegetation_cci project supplying continuous feedback, so these requirements will not be a surprise to the team.

Parameter	**Application**	**Horizontal Resolution**	**Observing Cycle**	**Length of Record**	**Precision**	**Accuracy**	**Stability**	Error Type (see Tab 3)	Source
LAI	GCOS 2022	10/100/250 m	1/5/10 d	>40/-/20 years	10%/-/20% for values >=0.5, 0.05/-/0.1 m2/m2 for smaller values	10%/-/20% for values >=0.5, 0.05/-/0.1 m2/m2 for smaller values	<3%/-/<6%		GCOS
	trend monitoring	< 1 km ~ 10 km	Daily ~ monthly	>30 years	0.05/-/0.1 m2/m2 (from low to high end)	10%/-/20% for values >=0.5, 0.05/-/0.1 m2/m2 for smaller values		ERRMERG	ECMWF
	model initialisation / boundary condition	< 1 km ~ 10 km	Daily ~ monthly	> 5 years	0.05/-/0.1 m2/m2 (from low to high end)	10%/-/20% for values >=0.5, 0.05/-/0.1 m2/m2 for smaller values		SSEOB/ERRMERG ERRCOV	ECMWF

	Global Reanalysis	< 1 km ~ 10 km	Daily ~ monthly	>40 years	0.05/-/0.1 m2/m2 (from low to high end)	10%/-/20% for values >=0.5, 0.05/-/0.1 m2/m2 for smaller values		SSECDR SSEOB/ERRMERG	ECMWF
	Monitoring/ Attribution	< 1 km ~ 10 km	Daily ~ monthly	> 5 years	0.05/-/0.1 m2/m2 (from low to high end)	10%/-/20% for values >=0.5, 0.05/-/0.1 m2/m2 for smaller values		SSECDR SSEOB/ERRMERG	ECMWF
	Data assimilation	< 1 km ~ 10 km	Daily ~ monthly	>5years ~NRT	0.05/-/0.1 m2/m2 (from low to high end)	10%/-/20% for values >=0.5, 0.05/-/0.1 m2/m2 for smaller values		SSECDR ERRMERG ERRCOV	ECMWF
	Model development and validation	< 1 km ~ 10 km	Daily ~ monthly	>5years ~NRT	0.05/-/0.1 m2/m2 (from low to high end)	10%/-/20% for values >=0.5, 0.05/-/0.1 m2/m2 for smaller values		SSEOB/ERRMERG	ECMWF
FAPAR	GCOS 2022	10/100/250 m	1/-/10 days	>40/-/20 years	5%/-/10% for values >=0.05, 0.0025/-/0.005 m2/m2 for smaller values	5%/-/10% for values >=0.05, 0.0025/-/0.005 m2/m2 for smaller values	<1.5/-/<3 %		GCOS
	Model development and validation	< 1 km ~ 10 km	Daily ~ monthly	>40~5 years	5%/for values >=0.05, and 0.0025/-/0.005 m2/m2 for smaller values	5%/for values >=0.05, and 0.0025/-/0.005 m2/m2 for smaller values		ERRMERG	ECMWF

Table 27: Summary of user requirements for vegetation ECVs.

4.26 Precursors for aerosols and ozone

Parameter	Application	Horizontal Resolution G/B/T	Observing cycle G/B/T	Uncertainty in measured value G/B/T	Stability G/B/T	Error Type (see Table 3)	Source
CO Tropospheric column	GCOS 2022	10/30/100 km	1/7/30 days	1/5/10 ppb	<1/1/2	2-sigma	GCOS
CO Mole							
HCHO Tropospheric column	GCOS 2022						GCOS

Table 28: Summary of user requirements for the precursors for aerosols and ozone ECVs.

4.27 River discharge

Climate change has a major impact on the Earth's water cycle, rivers are a significant component in this cycle and as such they must be monitored to understand the changes and interactions with other components. As well as the requirement to understand the effects of climate change on rivers in their own right, specific applications and areas of interest to climate research and services include the discharge of rivers into oceans which affects salinity and potentially ocean circulation and the impacts on human life and natural ecosystems which comes from extremes in river discharge, such as flooding and low river flow.

The current CCI project on river discharge is a precursors project to investigate the utility of satellite observations in providing long term monitoring of river discharge. It focuses on 15 river basins (see map at River Discharge ESA CCI project) for full list. Requirements gathering carried out by CMUG included

- comment: Which years/months are planned for these datasets?
- usage: this data can be exploited to compare with the simulated river discharge models, especially at the high frequency (daily and sub-daily).

Table X below lists the GCOS requirements for the River Discharge ECV and those obtained by CMUG from the Climate Modelling Community.

Parameter	**Application**	**Temporal resolution G/B/T**	**Observing cycle G/B/T**	**Uncertainty in measured value G/B/T**	**Stability G/B/T**	**Error Type (see Table 3)**	**Source**
River discharge	GCOS 2022	1/24/720 hours	1day/1 month/12 months	5/10/15 %	0.01/0.05/0.1 m/y/decade		GCOS

Table 29: Summary of user requirements for the river discharge ECV.

4.28 Other long lasting GHG

Parameter	**Application**	**Horizontal Resolution**	**Observing cycle**	**Error in measured value**	**Stability**	**Error Type (see Table 3)**	**Source**
	GCOS 2022						GCOS

Table 30: Summary of user requirements for other long lasting Greenhouse gas ECVs.

5. Across-ECV requirements

5.1 Common input/ancillary data, Links between ECVs and Dependencies

To ensure consistency between ECV datasets which is important for climate modelling and reanalyses there are a number of considerations that should be taken into account by all of the

CCI+ projects. To facilitate common practices the CCI should converge on terminology as this could potentially be different for each ECV project, agreed terminology will enhance communication across the CCI programme.

Firstly the ECV projects should all use the same level 1 datasets as input to their level 2 processing. Some of the ESA FCDRs (e.g. AATSR) were regenerated with improved calibration, geolocation etc. and guidance is needed from ESA, at least for ESA satellites, as to what are the recommended level 1&2 datasets to use. Table 26 shows which sensors are used by which ECV projects.

Secondly some ECVs will benefit from access to other ECV data sets from within the CCI programme to explore synergies and take advantage of opportunities where one ECV's retrieval can benefit from another. Table 27 attempts to identify where these cross-linkages are between ECVs.

Thirdly the use of common ancillary fields will be important. ERA5 should be the preferred source of atmospheric fields, which would ensure a consistent assumption about the atmospheric state for all ECV datasets. For surface fields an agreed SINGLE source for surface albedo, vegetation (LAI, FAPAR), emissivity, glacier climatology, sea ice, SST etc should be defined and agreed by the CCI projects. If this is not done inevitable inconsistencies will be seen in the products which will be only due to different representations of the atmosphere/surface being assumed. The common land/sea/lake mask produced by the LC CCI team also needs to be adopted by all ECV projects.

The horizontal grids should be common to level 3 products to enable easy comparisons and processing of data from different ECV CDRs. Similarly, the definition of atmospheric layering should be common across ECVs (e.g. aerosol and clouds) for level 2 and 3 products.

Finally, the specification of uncertainty characteristics should be provided in a consistent way and where appropriate separated into precision, accuracy and stability. The errors should also be specified, where possible, for each individual measurement.

Table 26. Primary sensors for each ECV project as given in last version of the DARD.

Inst / ECV	SST	Sea Level	Ocean Colour	Sea ice	Clouds	GHG	Aerosol	Ozone	Fire	HRLC	Land Cover	Glaciers	Soil Moisture	Greenland Ice Sheets	Antarctic Ice sheets	Water Vapour	Sea State	Biomass	LST	Perma frost	Lakes	Snow	SSS
AATSR/ATSR-1/ATSR-1	●				●		●												●	●	●	●	
ACE, SciSAT								●								●							
AIRS						●	●																
ALOS Palsar												●		●	●			●					
AltiKa																	●						
AMI-WS													●										
AMSR, AMSR-2/-E	●			●									●										●
AMSU						●																	
Aquarius																							●
ASAR G-POD											●												
ASAR, ASAR-WSM										●							●	●			●		
ASCAT													●										
ASTER												●		●	●								
AVHRR	●			●	●				●										●		●	●	
CALIPSO/CALIOP					●																		
CLOUDSAT/CPR					●																		
Cosmo-Skymed														●	●								
CRYOSAT 1/2		●		●										●	●								
CZCS			●																				
ESMR (Nimbus-5)				●																			
GEDI																		●					
GLAS																		●					
GOES IMAGER																		●					
GOME-1, 2							●	●															
GOMOS							●	●															
GOSAT						●																	
GRACE		●												●	●								
IASI							●	●											●				
ICESAT				●										●	●			●					
IRS1C/1D												●											
JAMI																			●				
JERS-SAR																		●					
Landsat TM/ETM+				●					●	●		●		●	●			●			●		
LDCM-OLI												●											
MERIS			●		●		●		●		●					●					●		
MIPAS								●								●							
MLS																●							
MODIS TERRA/AQUA			●	●	●				●										●	●	●	●	
OCO-2						●																	
OCM (OceanSat-2)			●																				
ODINO/SMR																							
OLCI					●				●							●					●		
OLI									●														
OMI							●	●															
OMPS								●															
OSIRIS								●															
Pleiades									●														
POLDER							●																
PROBA-V									●													●	
Radar Altimeters (ENVISAT, ERS, GDR, MGDR)		●		●										●	●		●	●			●		
Radar Altimeters (GEOSat Follow-on)		●																					
Radar Altimeters (JASON-1/2/3)		●															●				●		
Radar Altimeters (TOPEX-POSEIDON)		●																			●		
Radar Altimeters Follow on (GFO GDP NOAA)		●																					
Radar Altimeters Follow on CRYOSAT2		●		●													●						
Radar Altimeters Follow on Saral/Altika		●																			●		
Radar Altimeters Follow on Sentinel-3		●																			●		
RadarSAT														●	●								
SAR				●										●	●						●		
Scatterometers				●						●													
Sciamachy							●	●															
SeaWIFS			●																		●		
Sentinel-1 IWS																		●					
Sentinel-1 SAR									●	●							●	●			●		
Sentinel-2 MSI									●	●		●									●		
Sentinel-3 A/B																	●						
Sentinel-5-Precursor						●																	
SEVIRI	●				●		●												●				
SLSTR (Sentinel-3)					●				●										●			●	
SMAP																							●
SMMR				●									●										
SMR								●															
SMOS													●										●
SPOT										●	●											●	
SPOT VGT																							
SPOT-HRV												●											
SSMI-I & SSMIS				●									●						●				
TanDEM-X																		●					
TanSat						●																	
TerraSAR-X														●	●								
TMI	●												●										
TOMS								●															
VIIRS			●																				
WAVAS-II																●							
WINDSAT													●										●
WorldView										●													
Source DARD version or date web pages accessed	phase 2 D1.1	Issue 2.5 29/07/16	web: 09/10/20	web: 07/10/20	ATDB v5 12/09/17	web: 09/10/20	v3.3 02/05/17	v2.1 25/05/16	v2.5 30/11/17	v2.0 03/01/20	Web: 09/10/20	v1 20/22/11	web site accessed 09/10/20	Phase 2 D1.1	Phase 2 D1.1	v2 23/10/19	v1.1 27/03/19	v1 15/11/18	v1.1 17/06/19	v1.0 15/01/19	v1.1 05/09/19	Web: 09/10/20	v1.4 23/04/19

	SST	Sea level	Clds	Sea ice	OC	Aer'l	GHG	LC	Fire	O3	Glacs +ICs	Ice S.	SM	WV	SS Sal	SSt.	Lake	Snow	PF	LST	HR LC	AGB
SST		X	X	X	X	X								X	X	X				X		
Sea level	X			X				X			X	X	X		X	X						
Clouds	X			X	X	X	X	X	X	X			X	X	X			X	X		X	
Sea ice	X	X	X		X										X	X						
Ocean colour	X		X	X		X									X		X					
Aerosol	X		X	X	X		X	X	X	X			X					X				
GHG		X	X		X	X	X	X	X	X		X							X			
Landcover		X	X			X	X		X	X	X	X	X									
Fire			X			X	X	X		X	X		X									
Ozone	X		X	X		X	X															
Glacrs+ICs		X						X				X	X				X		X		X	
Ice Sheets	X	X			X						X									X		
Soil moisture		X	X				X	X	X		X											
WV	X		X												X					X		
SSS	X	X		X																		
S. State	X	X		X											X		X					
Lake											X							X	X	X	X	X
Snow																	X		X	X	X	X
Permafrost				X														X		X	X	
LST													X	X			X	X	X			X
HRLC								X									X	X	X	X		X
AGB								X	X				X					X	X			

Table 27: An analysis of cross linkages between ECVs indicating where links can be made towards consistency. The left hand column is the project with the ***identified need****, the top horizontal row is the* ***provider****. .*

5.2 Budget Closure Research

There have been three CCI related projects examining Earth system budgets of relevance to the climate system. These are the sea level budget closure project - CCI SL Budget Closure[20]; a carbon cycle budget project - RECCAP2: REgional Carbon Cycle Assessment and Processes[21]; and a mass balance project for ice sheets - IMBIE[22]. RECAPP2 and IMBIE are currently running, SLBC will restart for its next phase in late 2021. There is also a process-based project for understanding Arctic ocean processes – Arctic+[23]. All of these projects will benefit from spatially and temporally complete Earth observation datasets that are consistent and coincident (with respect to gridding and time step) with comprehensive descriptions of uncertainty from single observation to the full dataset, such as the CCI can provide.

Sea Level budget

Sea level is a crucial aspect of our ability to accurately monitor climate change. Essentially, the budget combines individual observation-based estimates of the net change in ocean mass due to the melting of ice sheets and glaciers, as well as the net change in ocean volume due to thermal expansion with an independent estimate of total sea-level rise from satellite altimetry. Note that on interannual timescales, accurate monitoring of the water cycle can also be important to track changes of water stored on land. The Sea Level Budget Closure[19] project aims to provide routine assessment of the global sea level budget, including estimates of the individual terms and their uncertainties. The requirements for sea level for such a study, defined by GCOS, are given in Table 30, and are consistent with the requirements gathered by CMUG and described in Section 4.14 for Sea Level Rise. The data outputs of the following CCI ECV projects are of interest to this budget project: SSH, SST, Glaciers, and Ice Sheets (Greenland and Antarctic).

Product	**Horizontal Resolution**	**Temporal Sampling**	**Required measurement uncertainty (precision and accuracy)**	**Stability**	**Source**
Global mean sea level	10-100km	Weekly to monthly	2-4mm (global mean); 1 cm over a grid mesh	< 0.3 mm/yr (global mean)	See EOV Specifications at www.goos ocean.org/eov
Regional mean sea level	10km	Hourly to weekly	1cm (over grid mesh of 50-100km)	< 1 mm/yr (for grid mesh of 50 – 100km)	www.goosocea n.org/eov

Table 28. Requirements for Sea Level as described in the GCOS in the Implementation Plan 2016 (GCOS-200) [Table 24].

---

20 https://tu-dresden.de/bu/umwelt/geo/ipg/gef/forschung/projekte/slbc_cci/sea-level-budget-closure-esa-cci-programme?set_language=en
21 http://cci.esa.int/reccap2
22 http://imbie.org/
23 https://arcticsalinity.argans.co.uk/

Carbon cycle processes and global Carbon budget

The data outputs of the following CCI ECV projects are of interest to the budget closure project on the carbon cycle (RECCAP2): AGB, GHGs, Permafrost, SST, and potentially OC.

6. Requirements for Climate Service Datasets

Climate data records (CDRs) are specialist data, defined by a US National Research Council committee, as "a time series of measurements of sufficient length, consistency and continuity to determine climate variability and change". In the US, agencies such as NASA and NOAA have sponsored operational and grant programs to create and curate CDRs. The focus has been on the reprocessing of satellite-derived records. The NOAA program includes a "research to operations" (R2O) initiative to transfer the development of routine and widely used data records from an individual PI's research group to an operational NOAA center.

In Europe, the European Space Agency (ESA) has launched the Climate Change Initiative to provide satellite-based CDRs (Hollmann *et al.*, 2013). The aspects related to R2O have been coordinated with the Copernicus Climate Change Service (C3S). Several datasets that were reprocessed under CCI Phase 1 are now available on the Climate Data Store (https://climate.copernicus.eu/climate-data-store).

KEY STRENGTHS:

- Many satellite records have been and are being reprocessed to correct for continuity problems arising from orbital drift, instrument degradation, replacing old satellites with new ones, improving and upgrading algorithms, etc.
- the CCI program includes efforts to characterize uncertainty in the records by means of running several retrieval products and comparing them with fiducial (ground-based) reference measurements
- The CCI/C3S data sets are created under a consistent definition and standards for data quality, documentation and automation

KEY LIMITATIONS:

- The CDR designation as used here may exclude many data sets that may be considered "climate quality" or otherwise suitable for climate research but are not part of the ESA or EU programs (or similar non-European initiative)
- CDR data sets are mainly satellite data sets and as such mostly do not cover periods longer than about 30 years (depending on the specific ECVs)

The main users of CDRs are in the climate research community. However, there has been more demand also from end-users, who require climate information in their decision-making. Policy makers, for example, are unlikely to use CDRs directly but require pre-digested information and indicators that are derived from CDRs. Some, such as risk modelling companies for the insurance industry, tend to be able to handle more 'raw' data, while other users, for instance farmers, require something processed to a format they can understand quickly. Translational users such as environmental consultancies, who sit in between the science specialists and the decision-maker, will add value to the data as part of a wider service offering. Some key areas here include:

- Sectoral studies – e.g. Agriculture: climate change impacts on yield, air quality impact on yield, pest and disease impacts
- Attribution studies (to link changes in datasets with man-made or natural events)
- Sustainability – climate change and impact on resources and environmental impact assessments (e.g. for large infrastructure projects)
- Resilience planning - climate change and impact companies and infrastructure, impact on vegetation (disease etc), urban planning/land use, insurance and reinsurance
- Climate fact sheets for regional studies or applications or to put in context current weather/climate events
- Adaptation studies and adaptation impact monitoring
- Hazard and event monitoring and information of the general public (e.g. drought events)
- “State of Climate” publications which summarize yearly the anomalies and trends of various ECVs either globally (https://www.ncdc.noaa.gov/bams/past-reports) or regionally (https://climate.copernicus.eu/european-state-of-the-climate#:~:text=The%20European%20State%20of%20the,behalf%20of%20the%20European%20Commission.)
- scientific publications and IPCC reports

For all of these applications, upper air and surface satellite data are important and will need to be combined with other available datasets. Typical requirements for datasets for *operational climate services* are:

- Simple user documentation on reading data and about the data characteristics
- Recognised format that is widely used (NetCDF4)
- DOI (from a recognized issuer)
- Uncertainty information on each parameter included
- Well validated as documented in a peer review paper
- Maturity matrix score documented and above a predefined value
- Ease of access on a recognized robust server with a given protocol (FTP)
- Timeliness for some ECVs (e.g. within 1 month of occurrence)
- Sustainability needing long term (>10yrs) archive commitment
- Ability for users to feedback comments on datasets to generators and other users
- Traceability
- Access to information of user applications
- Scientifically robust production e.g. through ensemble reforecasts or reanalyses (either based on ensemble methods, on variational approaches or on combinations of both)

Climate service users indicated a desire for three kinds of production chains:

a) Regularly updated (< 1 month) dataset with homogeneous processing that enables best comparison against long term record (note, this requirement can not be met by reanalysis).
b) Reprocessed climate records with new (improved) processing updated e.g. once per year. For example, improved retrieval algorithms that are applied to the same datasets when an upgrade is available.
c) Reprocessed climate records with more rigorous approaches such as variational or ensemble (or hybrid) reanalyses which are updated every 5 years and provide robust long-term data sets.

To address these needs C3S has designed the Climate Data Store (CDS), where users can find multiple datasets via a searchable catalogue for the applications of interest. The CDS provides a single point of access to a wide range of quality-assured climate datasets distributed in the cloud. CDS datasets include observations, historical climate data records, estimates of Essential Climate Variables (ECVs) derived from Earth observations (for example the ESA CCI datasets), global and regional climate reanalyses of past observations, seasonal forecasts and climate projections. Access to data is open, free and unrestricted.

Along with the data, the CDS includes a set of tools for analysing and predicting the impacts of climate change. Users of the CDS can access these tools to develop their own applications online. For example, the Climate Monitoring Facility developed in the first phase of CCI is now being integrated as an analysis tool in the CDS Toolbox, allowing a wider number of users an easier access to the data as well creating simple visualisations based on multiple data sources.

CDS data and tools form the backbone of the C3S Sectoral Information System (SIS), which provides tools and applications for dealing with climate impact in different industrial sectors, including energy, water management and agriculture.

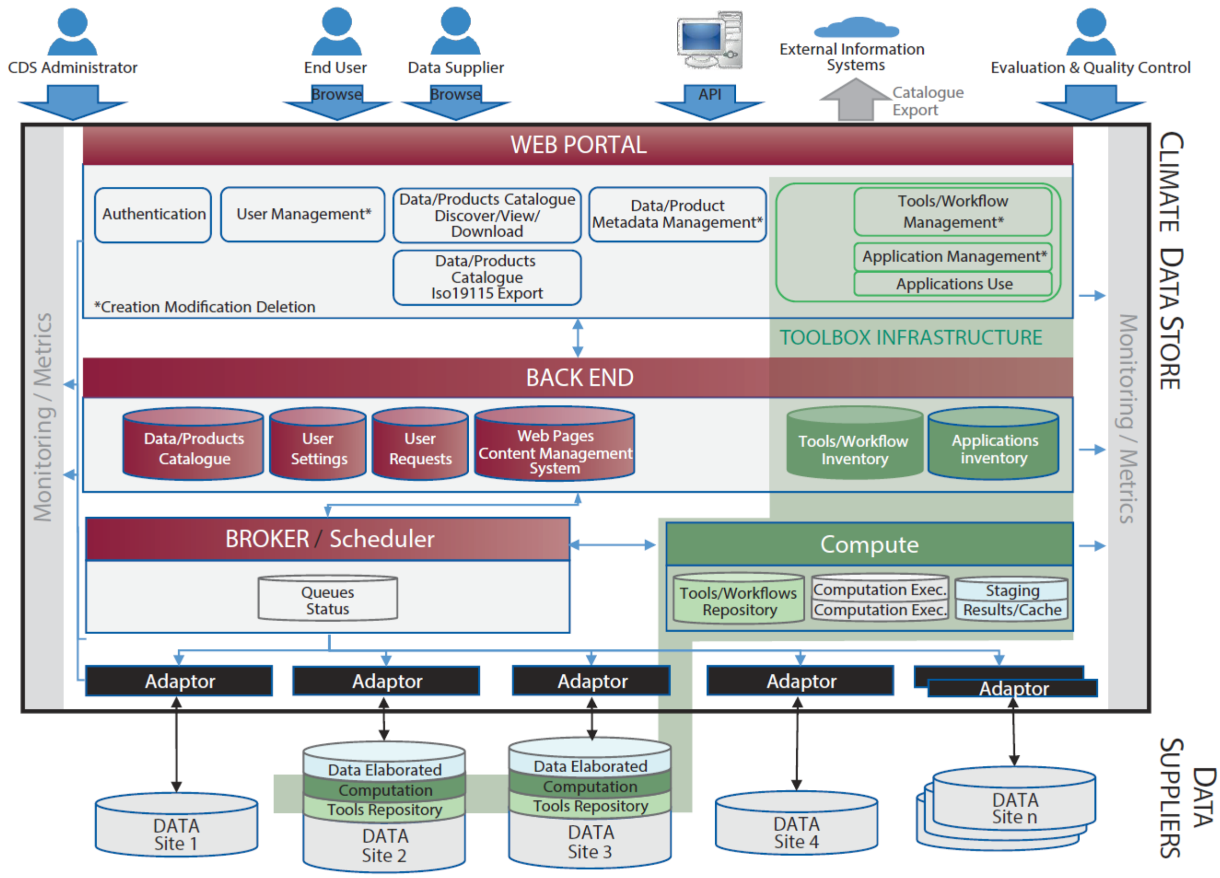


Figure X. The Climate Data Store infrastructure (source: https://climate.copernicus.eu/climate-data-store)

7. Requirements for obs4MIPs datasets

This section gives the current requirements that ECV datasets need to comply to in order to be included in the obs4MIPs database which is used by the CMIP modelling community for comparing satellite observations with climate model predictions (Teixeira *et. al*, 2014). All the information is also provided on the obs4MIPs site at: https://esgf-node.llnl.gov/projects/obs4mips/HowToContribute. We refer to this website for any updates on the requirements of obs4MIPs.

7.1 Criteria for Datasets to be included in Obs4MIPs

Observational datasets for obs4MIPs must fulfil the following criteria:

- Has clear traceability from level 1 measured variables to retrieved variables in level 3 or 4 dataset
- be based on data that has a history of peer reviewed publications,
- is version controlled, with doi,
- reside in a long term and maintained archive,
- span a time period long enough to be of use for model comparison (3 years is a useful minimum although in some circumstances shorter data records may be considered),
- match a model variable in the CMIP6 protocol
- include an estimate of the uncertainty for each variable verified by validation of the retrieved variables

7.2 Input Dataset Gridding

The datasets for consideration for obs4MIPs should be Level 3 (single sensor) or level 4 (multiple sensors) datasets which have been transformed on to a 1 degree grid square through averaging and/or interpolation and then averaged over 1 month. Researchers should be mindful to check on the obs4MIPs website that they are conforming to the latest specification. For each grid square the fields should be complete (i.e. no data voids) and consideration should be given to ensuring the variables are still conserved in the re-gridding. It is assumed only observational data (i.e. no model analyses) are included. The associated uncertainties also need to be provided on 1 deg grid and care has to be taken to derive these from the level 2 single field of view observations. Any biases in the original observations will propagate through to the gridded data but random errors will be reduced, hence averaging of uncertainties may not be appropriate. More details on the requirements are given here: https://esgf-node.llnl.gov/site_media/projects/obs4mips/ODSv2p1.pdf

The ECV teams are reminded that there are several issues to bear in mind when regridding data from level 2 to level 3 especially when dealing with uncertainties. These include maintaining consistency between variables after the regridding, consistently dealing with coastal areas correctly and how to fill data voids. It would be a good idea for data producers to share experiences on their regridding methodology.

7.3 Data Format

The file must be written in NetCDF version 4 and must follow the standard NetCDF Climate and Forecast (CF) Metadata convention https://cf-trac.llnl.gov/trac. The output file must pass a CF compliance check. A checker is available at: http://puma.nerc.ac.uk/cgi-bin/cf-checker.pl Choose the latest CF version when submitting the file for checking. Each output file must contain a time series of ONLY ONE physical variable (e.g. sea surface temperature, specific humidity). If the entire time series can be stored in less than 2GB, it must be stored in a SINGLE file. If it requires more than 2GB, it should be split into the minimum number of files required, with the size of each file being less than 2GB. Each file should contain a contiguous time series of complete data grid blocks. Each file must contain all of the required metadata applicable to the data subset contained in the file. Some support (tutorials, software) is provided at: https://pcmdi.llnl.gov/?cmip5/obs4cmip5.html.

Each physical variable and coordinate variable must use the specified output/coordinate variable name given in the CMIP5 Requested Output list (standard_output.xls). For example, the latitude output name must be “lat”, and the air temperature output variable name must be “ta”.

CCI+ CMUG WP5, led by DLR, is concerned with placing CCI ECVs into obs4MIPs and researchers on it will interact with the ECV teams to provide support in preparing their data for submission to obs4MIPs. The teams will also find it useful to share their experience on writing the compliant format datasets from the climate datasets, and feedback on any problems should also be given to the obs4MIPs team.

7.4 Documentation

A short technical note (5 pages max) must be provided with the dataset that conforms to the obs4MIPs technical note template: https://esgf-node.llnl.gov/site_media/projects/obs4mips/Obs4MIPs_Technical_Note_Guidance_v3.1.docx. It should be written bearing in mind the reader will not be familiar with satellite datasets. One important point to bear in mind is that there may be other datasets of the same variable available on the obs4MIPs site and so the note should make it clear what are the advantages of the CCI datasets with respect to previous datasets already available through obs4MIPs. These technical notes are valuable in their own right to promote the datasets and CCI teams should make them available on their web sites.

7.5 Process for submission of datasets

There is a proposal form for dataset owners to complete here: https://goo.gl/forms/GvhHmvYmMFQ_C210v2. The CCI project should keep a record of which datasets have been submitted to obs4MIPs and when they become available there.

8. Requirements for other ECVS

The climate data user community was asked to assess the priority of other ECVs for Climate Modelling and Analysis not covered by the ESA CCI. The results were that the following variables were of interest: Tropospheric nitrogen dioxide, seasonally frozen soil (not permafrost), near surface winds (10-100m above ground), surface solar radiation, and sea ice surface temperature. Also Fire radiative power as noted in section 4.21. Lastly one respondent raised the point about consistency between the products for snow albedo and snow, which they said has been lacking in the past.

9. Requirements for observation simulators

As well as satellite CDRs from the CCI projects for all 23 ECVs climate modellers also require observation operators, observation simulators or satellite simulators for some of the CDRs to convert the model state variables to the satellite measured variable are required. These operators are normally in the form of a generic software package that can be "plugged" into any climate model and interfaced with the model variables. The CFMIP Observation Simulator Package (COSP, Bodas-Salcedo *et. al.* 2011) is a good example of this and contains observation operators for many different satellite datasets, including Top of Atmosphere radiances, ISCCP, CloudSat, CALIPSO, HIRS and SSM/I.

The requirements for operators for each of the 23 ECVs will need to be considered. Currently it is envisaged that the observation operators listed in Table 29 will be required for the CCI datasets where the model variables are converted to a satellite observed quantity.

ECV	Model variable	Satellite variable to simulate
Atmospheric		
Cloud properties	Liquid/Ice concn profile	Cloud amount/top pressure
	Fractional cloud cover	Equivalent cloud cover
Ozone	Ozone concn profile	Total column ozone
Greenhouse gases	CO_2 and CH_4 profiles	Total column CO_2 and CH_4
Aerosols	Aerosol concn profile	Aerosol optical depth
Water vapour	Relative humidity	
Oceanic		
SST	Sea surface bulk temp	Sea surface skin temp
Sea level	Ocean dynamic topography	
Sea-ice	Sea-ice thickness	Area mean freeboard
	Sea-ice concentration	MW br. temps
Ocean colour	Phytoplankton concn	Derived chlorophyll alpha
Sea surface salinity	Salinity	Conductivity from microwave reflectivity
Sea state	Significant wave height	Ku band backscatter
Terrestrial		
Glaciers and ice caps	N/A	N/A
Land cover / HRLC	N/A	N/A
Fire / burnt area		
Ice sheets	Ice sheet thickness Ice sheet velocity	
Soil Moisture	Soil moisture	a) surface soil moisture b) surface saturation degree
Lakes	Lake surface bulk temp	Lake surface skin temp
Snow	surface_snow_thickness	
Permafrost		a) permafrost temperature b) active layer depth
LST	Surface bulk temperature	Surface skin temperature
AGB	Bio-stock volume	SAR backscatter / growing stock volume
Anthropogenic		
Total anthropogenic water use		

Table 29. Observation simulators required for CCI datasets.

The funding for the development and maintenance of an observation simulator package such as COSP is still not assured for many ECVs as it falls between the modelling community and observation community. The CCI project must ensure observations simulators for their observations are available to facilitate comparison with model fields.

10.Requirements for data formats and data access

The broad requirements for data formats and data access are outlined in the CCI Data Standards[24], below are any additional requirements which are not covered by this. In general the feedback from the CMC via the CMUG survey was that major obstacles in current satellite data usage are user friendly information and traceability, highlighting the need for better documentation and uncertainty information as the key issues lacking in current CCI standards.

10.1 Naming conventions and documentation

A data reference syntax is part of CMIP6[25], it outlines in detail the variables, units and naming conventions for each ECV. CMUG recommends that this document should be referenced in the next version of the CCI data standards document and where possible this syntax should be adopted for the CCI datasets also.

The short technical note for climate scientists with no knowledge of satellite datasets is recommended for each ECV. It would highlight the advantages of each datasets and its main characteristics. Specifically it should have an in depth description of error sources and uncertainties. The obs4MIPs guidance on technical notes[26] can be used for reference. Obs4MIPs remains the gold standard for quality in climate datasets, the guidance contains the following:

- *The target audience is the analysis community that will evaluate the climate model experiments in CMIP5, who have little experience with NASA datasets.*
- *The technical note should be written at the graduate student level.*
- *The note must be specific to one particular satellite observation dataset, which must contain a single variable.*
- *The note should summarize essential information for comparing the dataset to model output.*
- *Anything of interest only to experts should be referenced, but not include in the main body of the note.*
- *An appropriate length for the note (from Section 1 to 6 in the template) is 3-5 pages, excluding tables and figures.*

A technical note is mentioned in passing in the CCI data standards document, but CMUG recommends that more detail (as outlined above) is included.

10.2 Data formats

[24] https://climate.esa.int/sites/default/files/CCIDataStandards_v2-2_CCI-PRGM-EOPS-TN-13-0009.pdf
[25] http://proj.badc.rl.ac.uk/svn/exarch/CMIP6dreq/tags/latest/dreqPy/docs/CMIP6_MIP_tables.xlsx
[26] https://esgf-node.llnl.gov/site_media/projects/obs4mips/Obs4MIPs_Technical_Note_Guidance_v3.1.docx

The majority of CMUG survey respondents (83%) preferred NetCDF as a file format. This aligns nicely with the CCI Data Standards. For the metadata an XML document with a well-defined schema which clearly defines the instrument, its measurement technique and the analysis method used to retrieve the data record is recommended.

10.3 Data access

For getting access to the data 71% of the respondents requested FTP access, 37% requested web access via a browser (http:), 13% through OpenDAP, while some indicated a preference for access through another channel (NESDIS to NCEP, WMS, WCS, EUMETCAST, Wget, API). There is a need to be able to subset in time and space the datasets in a convenient way such as OpenDAP. Access from recognised data centres such as NASA DAAC, PMCDI and STFC were seen as a requirement reflecting the support they can provide to users.

In addition to access to the data from the CCI Open Data Portal, or from the Copernicus CDS (where some CCI data are already published) they could be hosted on a node of the Earth System Grid Federation (ESGF) so that users will have the same access interface for European, US and other climate datasets. They need to be hosted on the ESGF "data nodes" which publish to "gateway nodes", such as at BADC. As stated above obs4MIPs is the aspiration for all data types and CMUG recommends that work continues to add all ECV datasets the obs4MIPs database.

10.4 Level of processing

The user community was asked which level of processing they required for their applications from level 1 geophysical measurements (e.g. radiances), level 2 (derived products on original space view) or level 3 (e.g. daily, monthly means gridded products). The results are summarised in Table 32 which shows a fairly even split across the different processing levels.

Preference depended on the application. For assimilation, level 2 is required. For climate monitoring, level 3 is acceptable, but there must be traceability back to the sensor measurement and good documentation of the processing, because climate scientists need to understand how the variable has been calculated.

Processing Level	No. of users	Percentage of users
Level 1 + 2	4	8
Level 1 + 2 + 3	7	14
Level 2	6	12
Level 2 + 3	5	10
Level 2 + 3 + 4	6	12
Level 3	14	28
Level 3 + 4	7	14
Level 4	1	2
Total	50	100 %

Table 30. Feedback from users on required level of processing.

CMUG also sought user views on whether single sensor datasets or merged datasets would be required for level 3 gridded data products. The results depended on the ECV being considered. Table 33 suggests a preference for single sensor products, but the majority (57%) wanted both. Single sensor products are preferred by some for observation system simulation experiments (OSSE), bias correction etc. Some preferred merged products for better spatial and temporal coverage and more robust results, provided (again) that there is traceability back to the sensor measurement and good documentation of the processing. The disadvantage of merged products is that the error characteristics are more complex and single sensor products are preferred at level 1 or level 2 for reanalyses.

Single or Merged dataset	No. of users	Percentage of users
Single sensor datasets	13	28
Merged product datasets	7	15
Both	26	57
Total	46	100

Table 31. Feedback from users on single sensor vs merged products

Given this range of preferences from survey respondents, CMUG recommends that ECVs are made available at all possible processing levels and where possible both as merged and single sensor products.

10.5 Geospatial projections

Geospatial datasets have to be stored in a specific projection and this can cause problems in the analysis of the datasets (e.g. data day definition). The important thing is to provide simple tools to translate between any projection and a basic lat/lon grid. The CCI datasets should where possible share a common projection to facilitate the joint analysis of different datasets from different ECVs. Land/Sea/Lake and Cloud masks are also important to be common between the ECV projects, and those CCI masks produced in CCI Phases 1 and 2 should be propagated to the new ECV projects in CCI+ otherwise inconsistencies will be seen due to the use of different masks.

11. Summary

CMUG has carried out a survey of the requirements of the climate modelling community and present an analysis of these requirements here. One important finding is that the majority of modellers surveyed want to use the CCI datasets for model evaluation and development and only a few are engaged in climate monitoring.

An analysis of the individual requirements for climate modelling for the 23 CCI ECVs has been carried out with the following inputs:

- Inputs from CMUG interviews
- Comments and analysis through interaction by CMUG researchers with the climate modelling and reanalysis community and research meetings over the last year

This has enabled CMUG to undertake an analysis of how well the current GCOS requirements match the needs of climate modellers and how the ECV datasets in turn meet these needs. This information can be used as input to the CCI requirements specification as it evolves and is a good basis for discussions.

ESA CCI data should continue to be submitted to obs4MIPs to facilitate routine model evaluation with evaluation tools. The guidelines and specific requirements for obs4MIPs should be monitored and followed.

CMUG believes the CCI will meet the requirements listed here for most ECVs, and where it falls short this is due to limitations of the observational datasets. It is recognised that the climate observation data needs of the CMC can evolve, hence the need to re-consult at future dates with the CMC and revise this document accordingly.

The recent survey by CMUG for user requirements has shown that in many cases it is difficult for users to quantify uncertainty requirements for ECV data products. This is because there is a lack of quantitative information on the impact of different observation errors at different scales, for the range of applications addressed in this document. For critical applications, dedicated sensitivity studies are required to assess quantitatively the impact of uncertainties at specific spatial and temporal scales.

12.References

Banks, H., N. Bellouin, D. Bernie, S. Brown, N. Butchart, J. Caesar, D. Copsey, I. Culverwell, P. Falloon, P. Good, C. Harris, R. Hill, S. Ineson, T. Johns, C. Jones, A. Keen, J. Knight, S. Liddicoat, A. McLaren, A. Marshall, G. Martin, W. Moufouma-Okia, J. Rae, J. Ridley, J. Rodriguez, A. Scaife, A. Sellar, C. Senior, M. Vellinga, K. Williams, R. Wood, S. Woodward, and P. Wu, 2008: Evaluation of HadGEM3-AO. *Report to Defra March 2008.*

Bodas-Salcedo, A., M. J. Webb, S. Bony, H. Chepfer, J.-L. Dufresne, S. A. Klein, Y. Zhang, R. Marchand, J. M. Haynes, R. Pincus, and V. O. John, 2011: COSP: Satellite simulation software for model assessment. *Bull. Amer. Meteor. Soc.*, **92**, *1023–1043. doi: http://dx.doi.org/10.1175/2011BAMS2856.1*

Boucher, O., D. Randall, P. Artaxo, C. Bretherton, G. Feingold, P. Forster, V.-M. Kerminen, Y. Kondo, H. Liao, U. Lohmann, P. Rasch, S.K. Satheesh, S. Sherwood, B. Stevens and X.Y. Zhang, 2013: Clouds and Aerosols. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. *Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.*

Burke, E.J., Zhang, Y. and Krinner, G., 2020. Evaluating permafrost physics in the CMIP6 models and their sensitivity to climate change. The Cryosphere Discussions, pp.1-33. In press.

Eyring, V., Cionni, I., Bodeker, G. E., Charlton-Perez, A. J., Kinnison, D. E., Scinocca, J. F., Waugh, D. W., Akiyoshi, H., Bekki, S., Chipperfield, M. P., Dameris, M., Dhomse, S., Frith, S. M., Garny, H., Gettelman, A., Kubin, A., Langematz, U., Mancini, E., Marchand, M., Nakamura, T., Oman, L. D., Pawson, S., Pitari, G., Plummer, D. A., Rozanov, E., Shepherd, T. G., Shibata, K., Tian, W., Braesicke, P., Hardiman, S. C., Lamarque, J. F., Morgenstern, O., Pyle, J. A., Smale, D., and Yamashita, Y.: Multi-model assessment of stratospheric ozone return dates and ozone recovery in CCMVal-2 models, *Atmos. Chem. Phys.*, 10, 9451–9472, https://doi.org/10.5194/acp-10-9451-2010, 2010.

Gasser, T., Kechiar, M., Ciais, P. *et al.* Path-dependent reductions in CO_2 emission budgets caused by permafrost carbon release. *Nature Geosci* **11,** 830–835 (2018). https://doi.org/10.1038/s41561-018-0227-0

Giorgi, F, and Gutowski, W. J. 2016. Coordinated Experiments for Projections for Regional Climate Change. *Current Climate Change Reports* **2,** 202-210, https://doi.org/10.1007/s40641-016-0046-6

GCOS 2016: The Global Climate Observing System for Climate: Implementation Plan. *WMO GCOS-200 341 pp. Available at* https://library.wmo.int/doc_num.php?explnum_id=3417

GCOS 2022: The 2022 GCOS ECVs Requirements. *WMO, UNFCCC, GCOS-245. Available at* 202209121405---GCOS-245_2022_GCOS_ECVs_Requirements.pdf (unfccc.int)

permalink:
https://library.wmo.int/idurl/4/58111

GOFC-GOLD: 2016: Global Observation of Forest Cover - Global Observation of Land Dynamics - Annual Report, Wageningen University, Netherlands, 15pp. Published online at: http://www.gofcgold.wur.nl/documents/reports/gofc_po_apr2016.pdf

Hollmann, R., Merchant C. J., Saunders, R., Downy, C., Buchwitz, M., Cazenave, A., Chuvieco, E., Defourny, P., de Leeuw, G., Forsberg, R., Holzer-Popp, T., Paul, F., Sandven, S., Sathyendranath, S., van Roozendael, M., Wagner, W. (2013) The ESA Climate Change Initiative: Satellite Data Records for Essential Climate Variables *BAMS.*, **94**, 10, 1541-1552 https://doi.org/10.1175/BAMS-D-11-00254.1

Hyder, P., Edwards, J., Allan, R.P. *et al.* Critical Southern Ocean climate model biases traced to atmospheric model cloud errors. *Nat Commun* **9**, 3625 (2018). https://doi.org/10.1038/s41467-018-05634-2, https://www.nature.com/articles/s41467-018-05634-2

IGOS 2007: Integrated Global Observing Strategy Cryosphere theme report - For the monitoring of our environment from space and from Earth. Geneva: World Meteorological Organization. *WMO/TD-No. 1405, 100 pp.*

IPCC, 2013: AR5 WG1 Report, Climate Change 2013 The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. *(available at:* http://www.ipcc.ch/ipccreports/assessments-reports.htm).

IPCC, 2019: Summary for Policymakers. In: IPCC Special Report on the Ocean and Cryosphere in a Changing Climate[H.-O. Pörtner, D.C. Roberts, V. Masson-Delmotte, P. Zhai, M. Tignor, E. Poloczanska, K. Mintenbeck, A. Alegría, M. Nicolai, A. Okem, J. Petzold, B. Rama, N.M. Weyer (eds.)]. In press. https://www.ipcc.ch/site/assets/uploads/sites/3/2019/11/03_SROCC_SPM_FINAL.pdf

Lewis, E. L. & Perkin, R. G. (1978) Salinity: Its definition and calculation. *J. Geophys. Res. Oceans*, **83,** C1, 466-478. https://doi.org/10.1029/JC083iC01p00466

Loew, A., *et al*. (2017), Validation practices for satellite-based Earth observation data across communities, Rev. Geophys., 55, 779–817, doi:10.1002/2017RG000562.

Mapes, Brian & Chandra, Arunchandra & Kuang, Zhiming & Zuidema, Paquita. (2017). Importance Profiles for Water Vapor. 10.1007/978-3-319-77273-8_9.

Merchant, C.J., F. Paul, T, Popp, M. Ablain, S. Bontemps, P. Defourny, R. Hollmann, *et al*., 2017: Uncertainty information in climate data records from Earth observation. *Earth Syst. Sci. Data, 9, 511–527, 2017. https://doi.org/10.5194/essd-9-511-2017.*

Meehl, G. A., R. Moss, K. E. Taylor, V. Eyring, R. J. Stouffer, S. Bony and B. Stevens, Climate Model Intercomparisons: Preparing for the Next Phase, 2014: *Eos Trans. AGU, 95(9), 77, 2014.*

Scaife, A. A., Knight, J. R., Vallis, G. K. and Folland, C. K., 2005: A stratospheric influence on the winter NAO and North Atlantic surface climate. *Geophys. Res. Lett.*, **32**, *L18715, doi: 10.1029/2005GL023226.*

Seneviratne, S. I., Mueller, B., Koster, R. D., Orth, R. (2013) Role of soil moisture for (sub-)seasonal prediction. ECMWF Seminar on Seasonal Prediction, 3-7 September 2012 (https://www.ecmwf.int/sites/default/files/elibrary/2013/12178-role-land-surface-processes-seasonal-predictions.pdf accessed Oct 2020)

Seo, E., Lee, M., Jeong, J. *et al.* Impact of soil moisture initialization on boreal summer subseasonal forecasts: mid-latitude surface air temperature and heat wave events. Clim Dyn 52, 1695–1709 (2019). https://doi.org/10.1007/s00382-018-4221-4

Simmons, A. J., Poli, P., Dee, D. P., Berrisford, P., Hersbach, H., Kobayashi, S. and Peubey, C. (2014), Estimating low-frequency variability and trends in atmospheric temperature using ERA-Interim. *Q.J.R. Meteorol. Soc.*, **140**: *329–353. doi: 10.1002/qj.2317*

SPARC CCMVal (2010), SPARC Report on the Evaluation of Chemistry-Climate Models, V. Eyring, T. G. Shepherd, D. W. Waugh (Eds.), SPARC Report No. 5, WCRP-132, WMO/TD-No. 1526, http://www.atmosp.physics.utoronto.ca/SPARC.

Stubenrauch, C. J., *et al*., 2013: Assessment of global cloud datasets from satellites: Project and database initiated by the GEWEX Radiation Panel. Bull. *Am. Meteor. Soc.*, **94**, *1031–1049.*

Teixeira, J. D. Waliser, R. Ferraro, P. Gleckler, T. Lee, and G. Potter, 2014: Satellite Observations for CMIP5: The Genesis of Obs4MIPs. *Bull. Amer. Meteor. Soc.*, **95**, *1329–1334*. doi: http://dx.doi.org/10.1175/BAMS-D-12-00204.1

Toptunova, O., Choulga, M., and Kurzeneva, E.: Status and progress in global lake database developments, Adv. Sci. Res., 16, 57–61, https://doi.org/10.5194/asr-16-57-2019, 2019.

Turetsky, M.R., Abbott, B.W., Jones, M.C. *et al.* Carbon release through abrupt permafrost thaw. *Nat. Geosci.* **13,** 138–143 (2020). https://doi.org/10.1038/s41561-019-0526-0

Watson, A.J., Schuster, U., Shutler, J.D. *et al.* Revised estimates of ocean-atmosphere CO2 flux are consistent with ocean carbon inventory. *Nat Commun* **11**, 4422 (2020). https://doi.org/10.1038/s41467-020-18203-3

Zelinka, M. D., Myers, T. A., Mccoy, D. T., Po-Chedley, S., Caldwell, P. M., Ceppi, P., Klein, S. A., and Taylor, K. E.: Causes of Higher Climate Sensitivity in CMIP6 Models, Geophys Res Lett, 47, 2020.

13.Glossary

Terms	
Data assimilation	Observations directly influence the model initial state taking into account their error characteristics during every cycle of a model. This is used for reanalysis, NWP, which includes seasonal and decadal forecasting.
Model validation	Observations are compared with equivalent model fields to assess the accuracy of the model. This can be on short time scales for process studies or long time scales for climate trends.
Climate monitoring	This describes the use of a satellite only dataset to monitor a particular atmospheric or surface variable over a period > 15yrs to investigate whether there is a trend due to climate change.
Initialisation	To initialise prognostic quantities of the model with reasonable values at the beginning of the simulation but do not continuously update.
Prescribe boundary conditions	Prescribe boundary conditions for a model run for variables that are not prognostic (e.g. land cover, ice caps etc).
Accuracy	Accuracy is the measure of the non-random, systematic error, or bias, that defines the offset between the measured value and the true value that constitutes the SI absolute standard.
Stability	Stability is a term often invoked with respect to long-term records when no absolute standard is available to quantitatively establish the systematic error – the bias defining the time-dependent (or instrument-dependent) difference between the observed quantity and the true value.
Precision	Precision is the measure of reproducibility or repeatability of the measurement without reference to an international standard so that precision is a measure of the random and not the systematic error. Suitable averaging of the random error can improve the precision of the measurement but does not establish the systematic error of the observation.
Acronyms	
(A)ATSR	(Advanced) Along Track Scanning Radiometer on ERS -1&2 and ENVISAT
AGB	Above Ground Biomass
AVHRR	Advanced Very High Resolution Radiometer
BADC	British Atmospheric Data Centre
C3S	Copernicus Climate Change Service
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite
CAMS	Copernicus Atmosphere Monitoring Service
CCI	Climate Change Initiative
CCMVAL	Chemistry-Climate Model Validation Activity
CDR	Climate Data Record
CFMIP	Cloud Feedback Model Intercomparison Project
CMC	Climate Modelling Community
CMF	Climate Monitoring Facility
CMEMS	Copernicus Marine Environment Monitoring Service
CMIP5	Climate Model Intercomparison Project-5
CMUG	Climate Modelling Users Group
COSP	CFMIP Observation Simulator Package
CRDP	Climate Research Data Package

CSAB	Climate Scientific Advisory Board
DAAC	Distributed Active Archive Centres
DARD	Data Access Requirement Document
ECV	Essential Climate Variable
EGU	European Geophysical Union
ENSO	El Nino- Southern Oscillation
ERA	ECMWF Reanalysis
ERBS	Earth Radiation Budget Satellite
ERRMERG	Error of merged dataset
FAPAR	Fraction of Absorbed Photosynthetically Active Radiation
FCDR	Fundamental Climate Data Record
FOAM	The Fast Ocean Atmosphere Model
GCOS	Global Climate Observing System
GCW	Global Cryosphere Watch
GLDB	Global Lakes Data Base
GPS	Global Positioning System
GSICS	GCOS Satellite InterCalibration System
HIRS	High resolution Infrared Radiation Sounder
IGOS	Integrated Global Observing Strategy
IPCC	International Panel for Climate Change
ISCCP	International Satellite Cloud Climatology Project
LAI	Leaf Area Index
LST	Land Surface Temperature
MACC	Monitoring Atmospheric Composition and Climate
METAFOR	Common Metadata for Climate Modelling Digital Repositories
Met-UM	Met Office Unified Model
NAO	North Atlantic Oscillation
NWP	Numerical Weather Prediction
OSTIA	Operational Sea Surface Temperature and Sea Ice Analysis
PCMDI	Program for Climate Model Diagnosis and Intercomparison
PDO	Pacific Decadal Oscillation
PFT	Plant Function Types
RCM	Regional Climate Model
SAGE	Stratospheric Aerosol and Gas Experiment
SM	Soil Moisture
SSECDR	Single sensor uncertainty estimates for CDR
SSEOB	Single sensor error for each observation
SSH	Sea Surface Height
SSM/I	Special Sensor Microwave Imager

SSS	Sea Surface Salinity
SST	Sea Surface Temperature
TCDR	Thematic Climate Data Record
TEMPO	Tropospheric Emissions: Monitoring Pollution
UKESM1	United Kingdom Earth System Model version 1
UMARF	Unified Meteorological Archive and Retrieval Facility

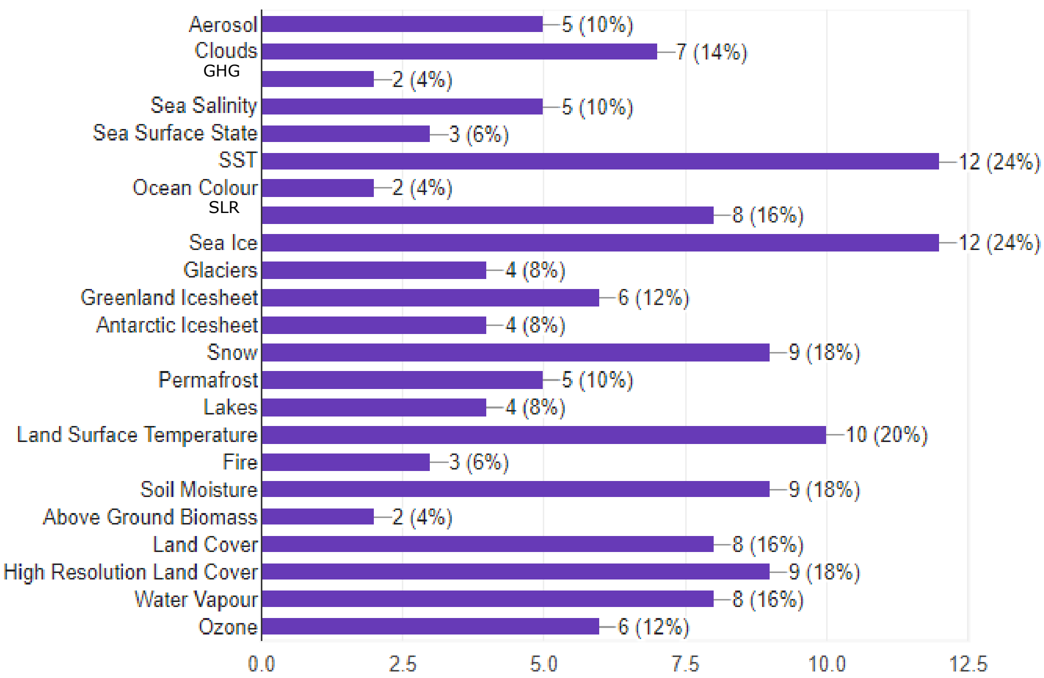
14. Annex 1 – Summary of expert views received from CMUG online survey

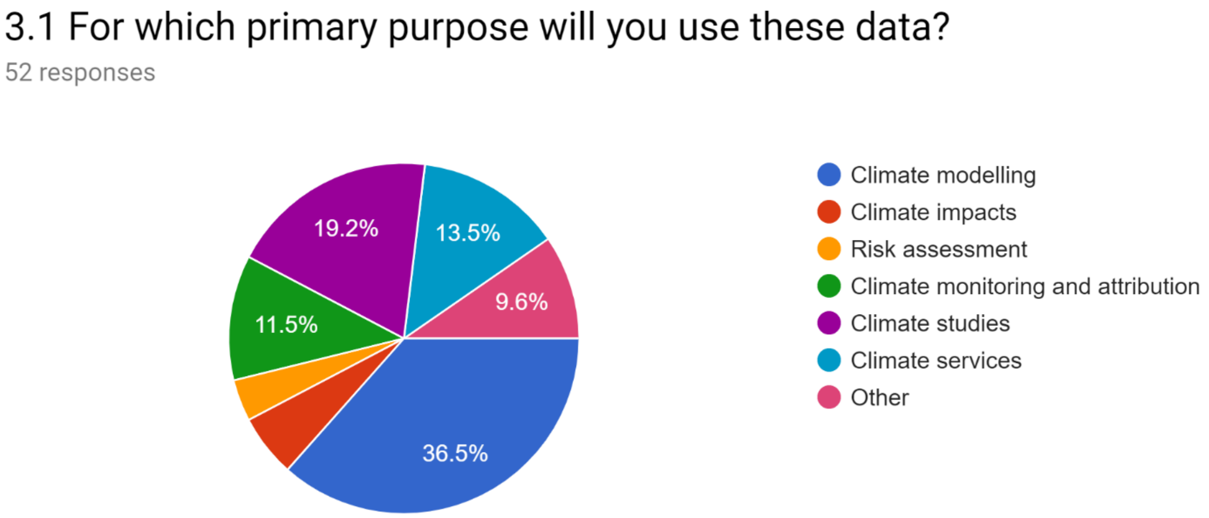
The contents of this annex are text inputs received from the experts who contributed to the CMUG online user requirements survey in early 2019, and the summary plots of the numeric responses received.

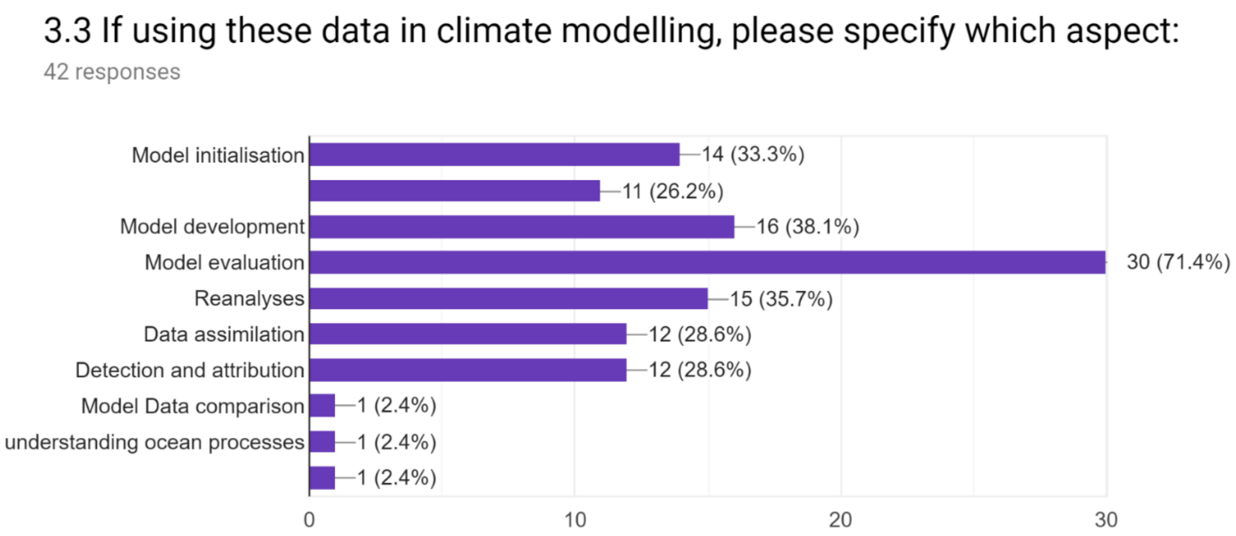
QUESTIONS RESPONSES 53

2.1 Please pick one of the following Essential Climate Variables (ECVs). If you wish to provide information about more ECVs, please complete a new survey each time.

50 responses

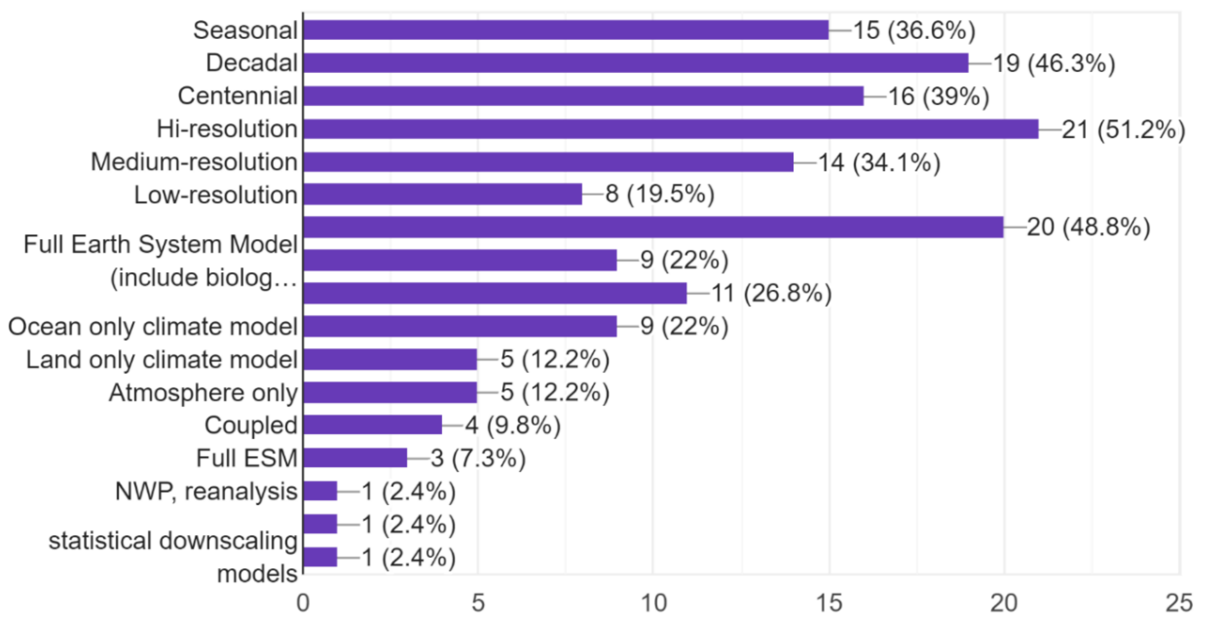



3.1 For which primary purpose will you use these data?
52 responses
19.2%
13.5%
9.6%
11.5%
36.5%
Climate modelling
Climate impacts
Risk assessment
Climate monitoring and attribution
Climate studies
Climate services
Other


3.3 If using these data in climate modelling, please specify which aspect:
42 responses
Model initialisation
14 (33.3%)
11 (26.2%)
Model development
16 (38.1%)
Model evaluation
30 (71.4%)
Reanalyses
15 (35.7%)
Data assimilation
12 (28.6%)
Detection and attribution
12 (28.6%)
Model Data comparison
1 (2.4%)
understanding ocean processes
1 (2.4%)
1 (2.4%)
0
10
20
30

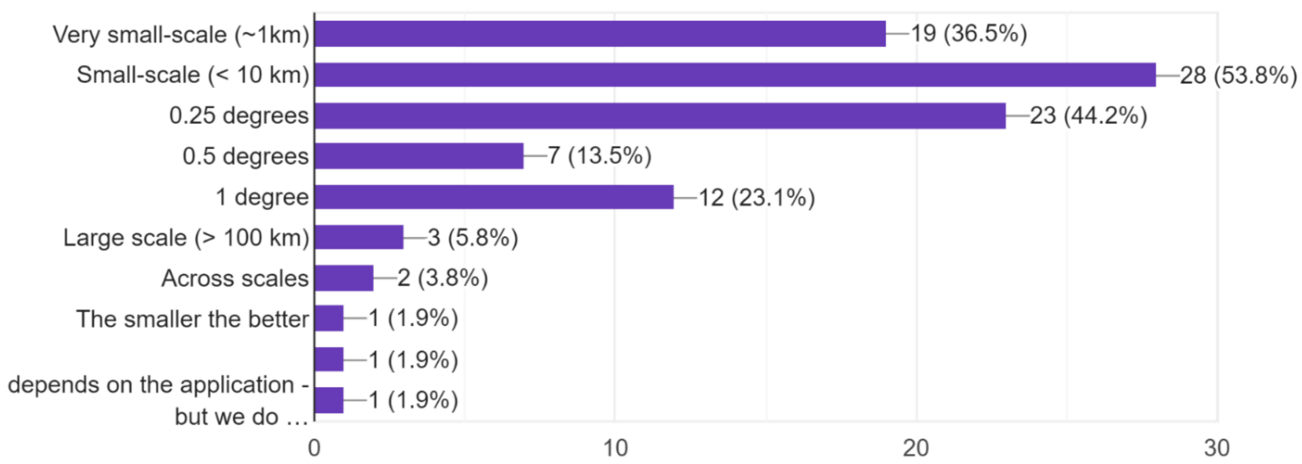
3.4 If you use these data in climate modelling, what type of modelling?

41 responses



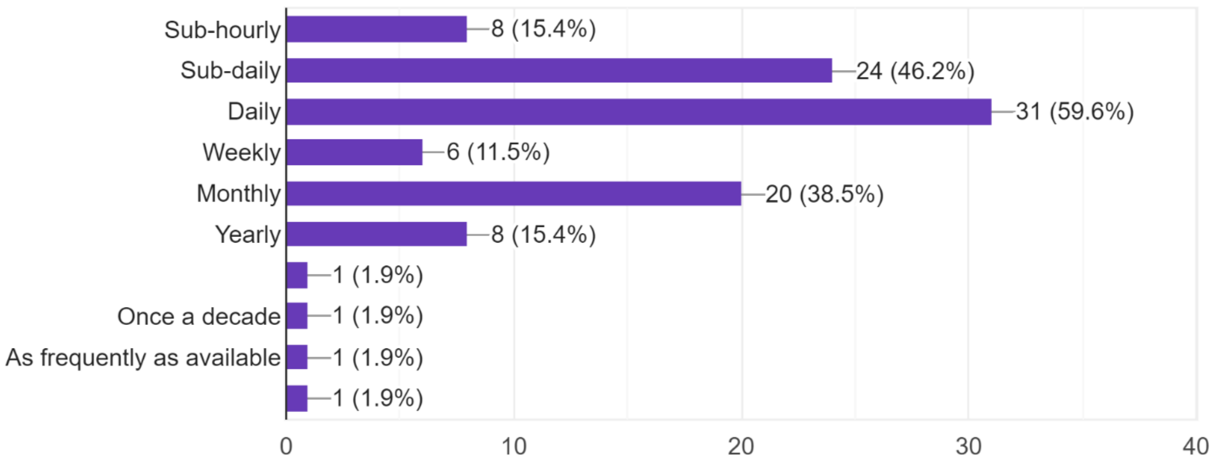
4.1 What spatial resolution(s) do you require of your ECV dataset of interest?

52 responses



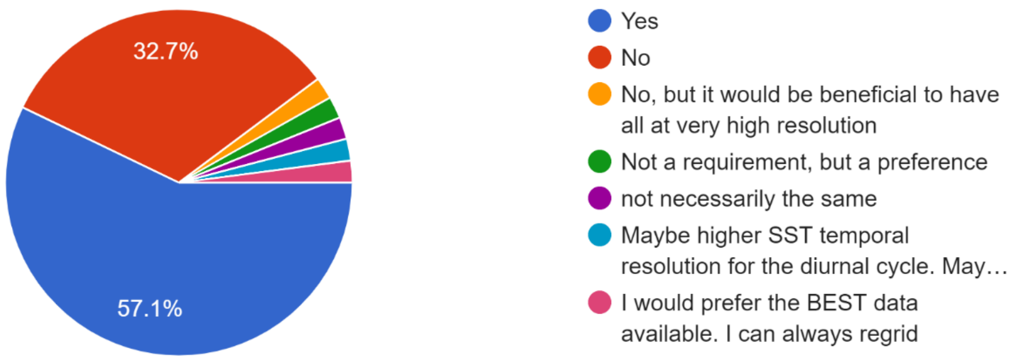
4.2 What temporal resolution(s) do you require of your ECV dataset of interest?

52 responses



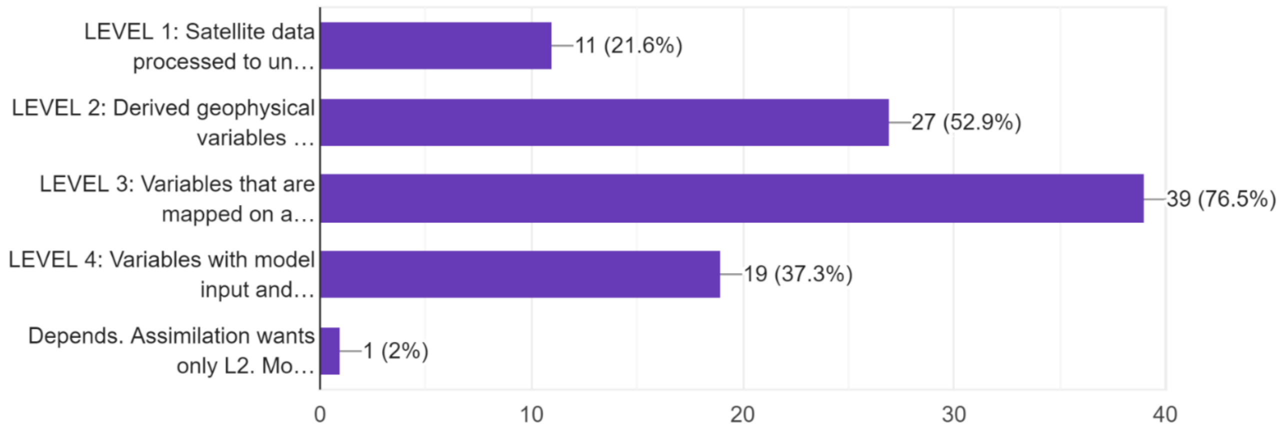
4.3 If you are using more than one ECV for your applications: do these products need a similar temporal/spatial resolution?

49 responses



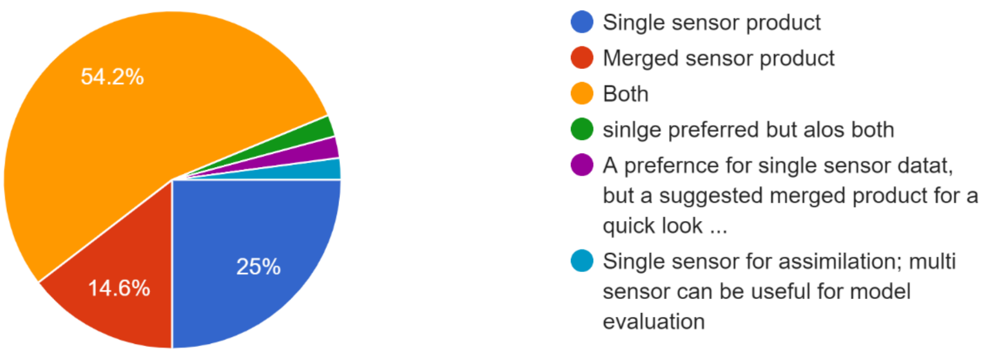
4.10 What Level of satellite observation data do you seek to use?

51 responses



4.11 Do you have a preference for single or merged sensor products, or both?

48 responses

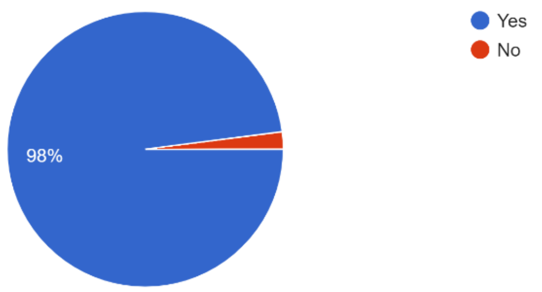


4.14 Are data gaps acceptable, (if flagged)?

52 responses

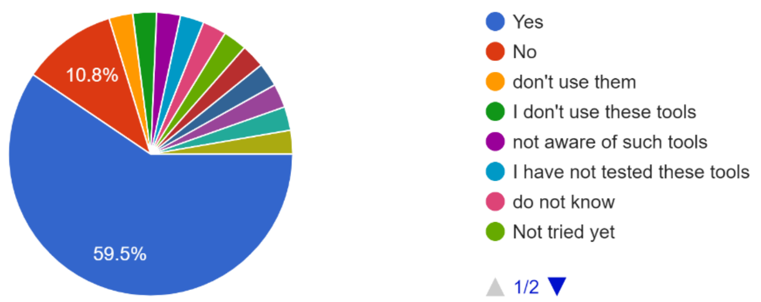
4.15 Are flags that describe a Quality Control issue desirable (e.g naming a source of noise)?

49 responses



7.7 Are the tools currently available for e.g. processing, visualising these ECV data satisfactory for your work?

37 responses



4.16 Do the GCOS requirements reflect your data needs? https://gcos.wmo.int/en/essential-climate-variables/requirements

42 responses